IN TWO SECTIONS—SECTION ONE

Transactions of the ASME

The Integration of Concepts in the Terminology of Measurement and Control . . *H. L. Mason* 1

Turbulence in a Tilting-Pad Thrust Bearing . . . . . *Stanley Abramovitz* 7

Investigation of Translatory Fluid Whirl in Vertical Machines . *G. F. Becker and B. Sternlicht* 13

Oil-Film Whirl—An Investigation of Disturbances Due to Oil Films in Journal Bearings. . . . . *B. L. Newkirk and J. F. Lewis* 21

Load Capacity and Time Relations for Squeeze Films . . . . . *F. R. Archibald* 29

Quadratic Programming of Interdependent Activities for Optimum Performance . *L. E. Saline* 37

Some Observations on Formal Models for Programming . . . . . *J. J. Slade, Jr.* 47

Theories on Bourdon Tubes . . . . . *F. B. Jennings* 55

Sensitivity and Life Data on Bourdon Tubes . . . . . *H. L. Mason* 65

Studies in Cold-Drawing—Part 3, Determination of Friction Coefficient . . . *H. Majors, Jr.* 79

Sulphur Dew-Point Corrosion in Exhaust Gases . . . . . *R. L. Coit* 89

Water Tunnels for Hydraulic Investigations . . . . . *J. M. Robertson* 95

An Extension of the Theory of Water Hammer. . . . . *R. Skalak* 105

Design of Optimum Clearances in Positive-Displacement Pumps and Motors . . *W. E. Wilson* 117

Fundamentals of the Vibratory Rate Gyro . . . . . *J. B. Chatterton* 125

Design Features, Operating Experience, and Performance of 100,000-KW Reheat Installation at Potomac River Station . . . . . *L. W. Cadwallader and H. S. Frederick* 127

An Approach to the Economic Problem of Matching Condenser Surface With Exhaust-Annulus Area . . . . . *W. A. Wilson and L. G. Maloof* 135

Improving Airplane Handling Characteristics With Automatic Controls . . . . *C. L. Maxsey* 143

Disturbance-Response Feedback—A New Control Concept . . . . . *J. B. Reswick* 153

Evaluation of the Nonlinear Kinetic Behavior of a Nuclear Power Reactor. . . . *R. C. Howard* 163

Aerothermodynamic Considerations Involved in Turbocharging Four and Two-Cycle Diesel Engines . . . . . *Rudolph Birmann* 171

Fuels for U. S. Navy Gas Turbines . . . . . *H. F. King and H. V. Nutt* 185

Effect of Size on the Inlet System Dynamics in Four-Stroke, Single-Cylinder Engines. . *D. H. Tsai* 197

Ultrasonic Testing of Small-Diameter Tubing With Automatic Recording Equipment. . . . . *W. L. Fleischmann and H. A. F. Rocha* 211

Heat Transfer to Turbulent Boundary Layers With Variable Free-Stream Velocity . . . . . *R. A. Seban and D. L. Doughty* 217

TRANSACTIONS OF THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS

VOLUME 78 JANUARY 1956 NUMBER 1

Transactions

of The American Society of Mechanical Engineers

Published on the tenth of every month, except March, June, September, and December

OFFICERS OF THE SOCIETY:

J. W. Barker, *President*

Joseph L. Kopf, *Treasurer* — C. B. Davies, *Secretary*

Edgar J. Kates, *Asst. Treasurer*

COMMITTEE ON PUBLICATIONS:

Otto de Lorenzi, *Chairman*

W. E. Reaser — John de S. Coutinho

Kerr Atkinson — B. G. A. Skrotzki

R. A. Crederero, H. N. Wendero } *Junior Advisory Members*

George A. Stetson, *Editor* — K. W. Clendinning, *Managing Editor*; J. A. North, *Asst. Managing Editor*

REGIONAL ADVISORY BOARD OF THE PUBLICATIONS COMMITTEE:

Roy L. Parelli—I
A. D. Blake—II
C. C. Franck—III
Francis C. Smith—IV
H. M. Cather—V
C. R. Earle—VI
M. B. Hogan—VII
Leon Helander—VIII

Published monthly by The American Society of Mechanical Engineers. Publication office at 20th and Northampton Streets, Easton, Pa. The editorial department is located at the headquarters of the Society, 29 West Thirty-Ninth Street, New York 18, N. Y. Cable address, "Dynamic," New York. Price $1.50 a copy, $12.00 annually for Transactions and the *Journal of Applied Mechanics;* to members $1.00 a copy, $6.00 annually. Add $1.50 for postage to all countries outside the United States, Canada, and Pan-American Union. Changes of address must be received at Society headquarters seven weeks before they are to be effective on the mailing list. Please send old as well as new address. . . . By-Law: The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B13, Par. 4). . . . Entered as second-class matter March 2, 1928, at the Post Office at Easton, Pa., under the Act of August 24, 1912. . . . Copyrighted, 1956, by The American Society of Mechanical Engineers. Reprints from this publication may be made on condition that full credit be given the Transactions of the ASME and the author, and that date of publication be stated.

The Integration of Concepts in the Terminology of Measurement and Control

By H. L. MASON,[1] WASHINGTON, D. C.

Problems in measurement and control arise in every field of science and technology, but many aspects are common or interdependent. Examination is here made of the origins, nature, spread, and fixity of the relevant American words and their glosses. Existing agencies are noted, and suggestions are made as to useful approaches to and procedures in terminology.

Introduction

THE central problem of terminology is the reduction of a mental image to language so phrased that it will re-create the original concept in the mind of a reader. Even in a field as broad as measurement and control, which touches the areas of mathematics, physics, chemistry, engineering, biology, and geology, there are concepts which unite diverse applications. It is the purpose of this paper to integrate these by examining some general principles by which the tangled threads might be woven into a coherent fabric.

What are the ends to be served? First of all, the provision for a meeting of minds in the discussion of new ideas. Such discussion, so heartily welcomed in professional gatherings, and so vital to exploration at the boundaries of knowledge, can probe the fallacies, establish the facts, convince the skeptics, reveal unsuspected applications. Although the scientist hears with scorn the demagogue who twists meanings, he sometimes forgets that truth can be effectively obscured by overspecialized jargon.

A second objective of lucid terminology is the guidance of the student who encounters a subject for the first time. His comprehension of new concepts and techniques depends on how well they are related to the sciences he has studied earlier. His sureness in applying them to the changing situations he is bound to meet in a lifetime of professional activity will be based on his over-all perspective of physical phenomena.

Lastly, precise language is necessary in the specifications needed for the design, construction, and test of equipment. The inventor's description of his patent, the engineer's instructions to his shop, the advertiser's appeal to his readers, the supplier's bid to his customer, the patent lawyer's handling of his clients' suits are all dependent for their effectiveness on the statement and the understanding of procedures and results expressed in technical terms.

Those pioneers who coined the word "acceleration" in 1531, or "thermometer" in 1633, reached only a few of their contemporaries with their treatises and journals in what was then called "Natural Philosophy." By 1953 the scientific and technical periodicals of the United States numbered over 4000, divided about equally between governmental and nongovernmental sources. Of these, about 1000 are in the basic sciences, and about 3000 in applied sciences such as medicine, agriculture, and engineering. Most of them devote some space to measurement and control in their special technologies, and there are a few (1)[2] which cover the field broadly. It is the contributors to these 4000 journals who develop new concepts and frequently introduce new words into the language. Often their works appear first as doctoral theses (2), or in research reports to government agencies (3). Within a year, perhaps, discussers at technical meetings and fellow workers in laboratories (the National Research Council listed about 1500 scientific and technical societies and associations as of 1948) have confirmed or extended the original usage. Before long the term appears in instrument textbooks (4), which nowadays aim to educate the practitioner as well as the college student.

The next step toward general acceptance of a word is the garnering and sifting procedure used in compiling technical dictionaries and encyclopedias, a score of which have been published in this country in the past decade (5). At this stage, classification of concepts are utilized by abstracting journals (6), library catalogs (7), and buyers' guides (8). Final acceptance may be considered to have been achieved when a term appears among 500,000 others in the general dictionaries, or in the Manual of Classification of the U. S. Patent Office.

Standardizing Agencies

The movement to give technical definitions a quasi-official status in the United States probably had its origin in the American Standards Association, organized in 1918 by five engineering societies. The ASA, which is "not a standardizing body, but a mechanism for achieving a consensus of opinion,"[3] delegates active sponsorship to one or another of its 102 constituent organizations. Among these are the American Institute of Electrical Engineers (9), The American Society of Mechanical Engineers (10), and the Institute of Radio Engineers (11), which with a few others (12) have been active in formulating the terminology of measurement and control.

In April, 1954, the ASA called a conference of 27 national groups to consider the procedures for standardizing terms dealing with automatic controls. It recommended the establishment of an ASA Sectional Committee to study the "Terminology pertaining to automatic process control, feedback control, regulating systems, and related systems not requiring human intervention as part of the regulating procedure."

The mere bulk of the job that needs doing is astounding. In the NBS Office of Basic Instrumentation, W. A. Wildhack, Chief, G. Keinath, and J. Stern have worked on an experimental

---

[1] General Engineer, Office of Basic Instrumentation, National Bureau of Standards. Mem. ASME.

Contributed by the Instruments and Regulators Division and presented at the Instruments and Regulators Division Conference, Philadelphia, Pa., September 13–24, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 7, 1954. Paper No. 54—IRD-3.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

[3] Address by C. T. Ainsworth, Technical Director.

system for indexing and locating instrument information. They compiled a list of 3500 nonproprietary names for instruments, 700 modifying adjectives, and 2700 conditions or quantities which can be measured. They accumulated a file of about 7000 definitions for these terms, as proposed by more than 60 professional or trade groups. There are many instances of different terms for identical concepts, and of conflicting definitions of the same word.

Let us consider some of the basic principles which a terminology committee should keep in mind.

Survey by Concepts

In 1852 Peter Mark Roget published a Thesaurus which has since run to an estimated ten million copies. He was a physician by training, but his intellectual achievements included a professorship of physiology, a study of Metropolitan London's water supply, the invention of a slide rule, the construction of a calculating machine, the writing of technical papers on electricity, galvanism, and magnetism, and a 20-year stint as secretary of the Royal Society. With this breadth of interest, it is not surprising that his grouping of the concepts expressible in the English language is considered the most comprehensive ever made. The main categories of interest to us are: (I) Abstract Relation, including Existence, Kind of Relation, Quantity, Order, Number, Time, Change, and Causation; (II) Spatial Relations, including Dimensions, Form, and Motion; (III) Matter.

It is suggested that the committee use Roget's subheads as a kind of check list, both for estimating the completeness of its work, and for comparing the results with those of other groups. For example, it would find under subhead 12 several ways of expressing "Correlation," by such terms as "corresponding," "reciprocal," "interconnected"; 27, "Equality," by "symmetry," "equilibrate," "isochronous"; 62, "Precedence," by "leading," "foregoing," "in advance." An approach of this sort is indicated by ASA Z58.1.1, which lists Processes, Devices, Quantities, and Properties of Devices. Again, Committee C42 of ASA uses a very detailed breakdown in its "Definitions of Electrical Terms," with about 3800 items. There is one group of general terms, and 18 other groups specific to classes of equipment or field of use. Each group has 5 to 20 sections, either detailing subclasses of equipment, or bearing such headings as Kind of Protection, Qualifying Terms, Properties and Characteristics, Design Details, Duty and Services, Testing, General, Not Otherwise Classified. The reader finds his way into the structure via an alphabetical index.

Universality of Concepts

It may be taken as axiomatic that there are many notions which are common to several fields of science or technology. For example, dy/dx represents a mathematical symbol, the slope of a line, the rate of change of the physical variable y, or the ratio between applied emf and scale reading of a voltmeter. Again, "that which has the ability to influence" might serve as a generic definition of force—in mechanics, electric circuits, magnetism, acoustics, or as the right-hand member of a differential equation in its standard form.

A committee can extend the range and prolong the life of its work by phrasing its definitions in the broadest possible terms. Inclusion in its membership of representatives from distinct fields of application is one technique for accomplishing this. Another is the circulation of a proposed standard among other professional groups—as many as 30, in the case of the Applied Spectroscopy Committee of SAS-ASTM. This group notes that conventions regarding suffixes are now fairly well established: -or, a device; -ation, a process or result of a process; -ance, a property of a device or body; -ity, property of a substance; -meter, a measuring device; -scope, a viewing device; -graph, a recording device; -gram, the record produced; -scopy, observation. They note that for the suffixed -or, -ation, -ance, -ity, a tabulation of 544 entries from eight languages shows that one or another is used in 96 per cent of the total.

As regards undesirable usage, the SAS-ASTM group recommends that if a term is faulty in some important respect, an effort should be made to change it; if only slightly faulty but too firmly established to be discarded, a brief explanatory comment should be included in its definition. On the other hand, the AIEE-IRE-ISA Committee on Telemetering has the basic rule of using existing definitions wherever possible. One wonders why this was not invoked to avoid undue restriction in their definition of "Detector: A device which performs the demodulation of an a-m r-f signal to a-f," or again in their definition "Remote Control: A system . . . (using) electric means over one or more conductors per unit . . . ."

Exactness of Language

A cardinal principle of the definition writer should be to make sure that his words do not contravene more general usage in expressing his own specialized meaning. Now that we have power generated by bits of germanium exposed to light, is it still feasible to use the term "thermoelectric power" to express millivolts per degree for a thermocouple? What is the student of servomechanisms to think when he reads that "velocity error" is measured in units of position (and is positive when smaller than a reference value), or finds that a "step response" seldom takes the form of a step? How many engineering writers mistreat mathematics, using "inverse" when they mean "of opposite sign," or "term" when they mean "factor"?

Let the committee include in its membership at least one professional writer, or perhaps a college professor who can tap the critical resources of the Department of English. This may help it steer a middle course between the precise but wordy patent claim which describes ". . . electroresponsive means operatively associated with said movable element and adapted to be influenced by its frequency of vibration. . .," and the string of qualifying nouns common to stockroom catalogs which list a "Wrench, starboard aft gudgeon cap stud nut socket."

Style

Perhaps a pleasing literary style is something that is born, not made, but a good lexicographer has some tricks that are worth noting. He uses as many words as needed, but no more. His sentence structure fits the logical relationship of his ideas: if he is discussing initial and final members, he mentions the former first and the latter last. He puts the word "only" next to the noun or verb it modifies. He hangs dependent clauses at the point where they apply, not half a sentence away. In discussing successive aspects of a situation, he prefers to repeat key words rather than to use synonyms which might add color but lessen clarity. His sentences tend to be short and his vocabulary simple. The creed of ASA Committee C42 is a good foundation: "To express for each term the meaning generally associated with it in electrical engineering work in the United States, tending toward a simple statement of function, and generalized where possible so as not to preclude specific interpretations."

Orthographic Projections

Every engineer is familiar with the representation of an object by means of its projections on three or more planes. Technical concepts might well be examined in similar fashion, turned over in the mind with the hope that, when viewed from the same angle, their similarities will be evident.

Perhaps the commonest viewpoint is that of appearance; an

entity is named in accordance with its shape, or size, or structure. Thus we have the plane table, the coil, the coarse adjustment, the megawatt, the beam of electrons. Closely akin to this is a description, not of the object itself, but of some useful representation. Thus we have linear equations, waves of pressure, white noise.

Another viewpoint is that of function. From this angle we observe that nearly every measurement involves something which may be called a detector, and something else which functions as a standard. Then there are resistors, microscopes, signals, projectors, isolating amplifiers, sinks, controllers and controlled systems (although some prefer the ambiguous "control element"). The SAE requires that "names shall be functional, in that they describe the service the part performs, the purpose for which it is used, or its location in an assembled unit." The SAS-ASTM group recommends that instruments should be named for the quantity they measure.

Another viewpoint is that of origin. Forces which are called dynamic arise from change. The name of the element californium shows geographical influence; the gas helium was first noted in the sun. Original investigators are brought to mind by Young's modulus, the Geiger counter, the Nyquist diagram.

Also very common are designations which describe the way a thing behaves. The slow neutron is tardy, diffraction breaks up light, viscosity implies stickiness.

Recognition of such orthographic views of a concept is one way of unifying the terminology of independent fields. Agreement on what is the most important aspect of a concept can lead directly to agreement on the name to be given it. For example, the current controversy on "automatic-control system" versus "feedback-control system" could be resolved into the question of whether it is preferable to emphasize the over-all purpose or the manner of its accomplishment.

Preferred Terms

The compiler of a dictionary usually considers himself a reporter rather than an inventor or a critic. Typical prefaces state that he "must above all record usage"; that it is his "function to inform by accurate description"; that he must realize that "many technical terms are misnomers." Knowing that usage is frequently specialized and sometimes contradictory, he does not hesitate to set forth multiple meanings, prefacing each with a restrictive phrase: "In science generally," "In aeronautics," "In a nuclear reactor," "In petroleum technology." One recent dictionary goes somewhat further, applying what it calls "the keyword plan: Each topic is defined in basic terms, and then every word that is further defined elsewhere in the book is printed in bold-face type."

A terminology committee often works in fields where no consensus exists. It has the privilege and the duty of exercising its best judgment in selecting a preferred term from among several synonyms that may have been used to denote a given concept. For example, in control terminology, it might prefer the word "transference," rather than transfer function, transmittance, admittance, adpedance, or frequency-response function. It might borrow from other usages such terms as signal, response, comparator, corregidor, actuator. It might coin new words from Greek or Latin roots, as has been done in cybernetics, automation, nucleonics. It might even adopt a practice which spread rapidly among Federal agencies in the 1930's, and contract a lengthy name into a set of initials. We have long recognized d-c and a-c, and now come the convenient rf, uhf, AVC, and ILS.

Alternative Usages

What is a terminology committee to do when it is presented with equally sound arguments supporting two or more names for the same concept? It might agree to disagree, and omit the concept altogether from its listing, but that is outright dereliction of duty. A better way is to adopt the dictionary's practice of listing each of the controversial terms in its usual alphabetical position, but referencing it to the preferred term, where there appear a set of near-synonyms, glossed to indicate the circumstances appropriate for their use. Usages which are obsolescent, or deprecated, or outright erroneous can readily be indicated in this fashion, with great convenience to the user. Thus the definition of "error" should show its relation to accuracy, correction, and deviation. The word "sensitivity" deserves similar treatment.

Again, multiple definitions are sometimes needed, as for example: "Accuracy: (1) Of a single measurement, its closeness to an accepted standard value, expressed as their difference (its error). The property is the joint effect of method, observer, instrument, and environment. It is impaired by mistakes, by random error, and by bias. (2) Of a repetitive set of measurements, the closeness of their mean to an accepted standard value, expressed as their difference. (3) Of a measurement not yet made, the deviation (from an accepted standard) not likely to be exceeded (say 99.7 per cent of the time). It should not be attributed solely to method, or observer, or instrument, or environment unless the other three factors of the four are standardized. (4) Of an instrument, the closeness with which its readings agree, or may be expected to agree, with specified standards under specified circumstances; accuracy rating. (5) Of control systems, the closeness with which the instantaneous (or the finally asymptotic) value of the controlled variable agrees with the value desired."

I might add that Committee C39.2 of ASA, in trying to correlate the views of various groups on the word "accuracy," has assembled a six-page monograph on the subject.

Arrangement

Committees developing definitions for a narrow field, say, automatic substation equipments, may find twenty or thirty terms sufficient for their needs. In that case, arrangement by concepts seems feasible for the final publication. It might group terms relating to major elements, those relating to performance, and those relating to parts. From the user's viewpoint, this is most desirable for integrating the concepts of one field with those of another.

This scheme becomes unwieldy when totals run into the hundreds. Then the usual dictionary arrangement seems best, allowing the user to locate a word and its definition without recourse to an index. Such a scheme was adopted by the AIEE Committee on Feedback-Control Systems in their proposal of April, 1954. The IRE currently issues a mimeographed alphabetical list of some 6000 terms on which its committees have worked or are working, and references each to a published standard or active committee from which the definition may be obtained. This compilation appears to be a very useful device; it provides for continual change, keeps committees informed of each other's work, and permits early publication of small glossaries for narrow fields.

Conclusion

The foregoing observations have indicated the state of the art in American terminology for measurement and control, and suggested some specific approaches to its problems. If the reader is stimulated to lend a hand to the integration of concepts, my purpose will have been accomplished. He will need the precision of a Kelvin, the imagination of a da Vinci, and the wisdom of a Solomon—but let him not be deterred!

BIBLIOGRAPHY

1 "Automatic Control," Reinhold Publishing Corp., New York, N. Y.

"Automation," Penton Publishing Co., New York, N. Y.

"Computers and Automation," by E. C. Berkeley Associates, New York, N. Y.

"Control Engineering," McGraw-Hill Book Company, Inc., New York, N. Y.

"Industrial and Engineering Chemistry," American Chemical Society.

"Instruments and Automation," Instruments Publishing Company.

Journal of the Instrument Society of America.

Journal of Applied Physics, American Institute of Physics.

Journal, Association for Computing Machinery.

Proceedings, Institute of Radio Engineers and its professional groups.

Proceedings, Eastern Joint Computer Conference; Western Joint Computer Conference.

"Review of Scientific Instruments," American Institute of Physics.

Transactions, AIEE.

Transactions, ASME.

2 "Guide to Bibliographies of Theses," by T. R. Palfrey and H. E. Coleman, second edition, American Library Association, 1940.

"Compilation of PhD and Master's Theses in Chemistry and Chemical Engineering," an annual publication by the American Chemical Society.

"Dissertation Abstracts," by University Microfilms, Ann Arbor, Mich.

3 "Bibliography of Technical Reports," by the Office of Technical Services, Washington, D. C., monthly, 1946—.

4 "Instrumentation Literature and Its Use," by J. F. Smith, Office of Basic Instrumentation, 1952, lists some 150 textbooks and reference works published in the United States, covering the measurement of dimensions, weight, position, geodetics, photogrammetry, heat, light, sound, ultrasonics, electricity, electronics, radio, nuclear energy, pressure, vacuum flow, stress; also theory and design of instruments, servomechanisms, and process controls.

5 "ASTM Glossary of Terms," Committee E8, Philadelphia, Pa., 1951.

"Concise Chemical and Technical Dictionary," by H. Bennett, Chemical Publishing Company, 1947.

"Engineering Terminology," by V. J. Brown and D. G. Runner, second edition, Gillette Publishing Company, 1939.

"Callaway Textile Dictionary," by W. L. Carmichael, et al., Callaway Mills, 1947.

"Chemical and Engineering Dictionary," Chemical Publishing Company, 1944.

"Electronics Dictionary," by N. M. Cooke, McGraw-Hill Book Company, Inc., New York, N. Y., 1945.

"The Practical and Technical Encyclopedia," by J. A. Damm and C. H. Waugaman, W. H. Wise & Co., New York, N. Y., 1948.

"A Dictionary of 10,000 Scientific and Technical Words," by W. E. Flood and M. P. West, Longmans, Green & Co., New York, N. Y., 1952.

"Baughman's Aviation Dictionary," by E. J. Gentle and C. E. Chapel, third edition, Aero Publishing, Inc., Glendale, Calif., 1951.

"Dictionary of Applied Physics," by R. T. Glazebrook, The Macmillan Company, New York, N. Y., 1922.

"Hackh's Chemical Dictionary," by J. Grant and I. W. Dagobert, third edition, The Blakiston Co., New York, N. Y., 1944.

"Audels' Mechanical Dictionary," by N. Hawkins, Audel and Company, 1942.

"Metallurgical Dictionary," by J. G. Henderson, Reinhold Publishing Company, 1953.

"Dictionary of Mathematical Sciences," by L. J. Herland, F. Ungar Publishing Co., New York, N. Y., 1951.

"Bibliography of Interlingual Scientific and Technical Dictionaries," by J. E. Holmstrom, second edition, UNESCO, Paris, France, 1951.

"Mathematics Dictionary," by Glenn James and R. C. James, D. Van Nostrand Company, Inc., 1949.

"Dictionary of Conformal Representations," by H. Kober, Dover Publications, New York, N. Y., 1952.

"Dictionary of Statistical Terms and Symbols," by A. K. Kurtz, Cherrydale, Va., 1935.

"A Dictionary of Color," by A. J. Maerz, second edition, McGraw-Hill Book Company, Inc., New York, N. Y., 1950.

"Metals and Alloys Dictionary," by M. Merlub-Sobel, Chemical Publishing Company, 1944.

"A Dictionary of Metallurgy," by A. D. Merriman and J. S. Bowden, *Metal Treatment*, vol. 20, 1953, pp. 604–610.

"Illustrated Technical Dictionary," by M. Newmark, Philosophical Library, 1944.

"Blakiston's New Gould Medical Dictionary," The Blakiston Co., New York, N. Y., 1951.

"Encyclopedia on Cathode-Ray Oscilloscopes," by J. F. Rider and S. D. Uslan, published by J. F. Rider.

"Dictionary of Industrial Engineering," by D. S. Shybekay, W. Lafayette, Indiana, 1945.

"Scientific Dictionary," by J. R. Stewart, fourth edition, Stewart Research Laboratory, 1953.

"Chambers' Technical Dictionary," by C. F. Tweney and L. E. C. Hughes, The Macmillian Company, New York, N. Y., 1949.

"Hutchinson's Technical and Scientific Encyclopedia," by C. F. Tweney and I. P. Shirshov, The Macmillan Company, New York, N. Y., 1936.

"Chemist's Dictionary," D. Van Nostrand Company, Inc., New York, N. Y., 1953.

"Scientific Encyclopedia," second edition, D. Van Nostrand Company, Inc., New York, N. Y., 1947.

"A Dictionary of Electronic Terms: Radio, Television and Electronics," by H. L. Van Velzer, Allied Radio Corporation, 1950.

"Glossary of Physics," by L. D. Weld, McGraw-Hill Book Company, Inc., New York, N. Y., 1937.

6 Abstracting Journals:

"Aeronautical Engineering Review," Institute of Aeronautical Sciences.

"Applied Mechanics Reviews," ASME.

"Battelle Technical Review," Battelle Memorial Institute.

"Biological Abstracts," University of Pennsylvania.

"Ceramic Abstracts," American Ceramic Society.

"Chemical Abstracts," American Chemical Society.

"Engineering Index," Engineering Societies Library.

"Gas Abstracts," Institute of Gas Technology.

"Geophysical Abstracts," U. S. Geological Survey.

"Mathematical Reviews," American Mathematical Society.

"Meteorological Abstracts," American Meteorological Society.

"Nuclear Science Abstracts," U. S. Atomic Energy Commission.

"Physics Abstracts," (British) Institution of Electrical Engineers, American Institute of Physics.

Proceedings IRE: Abstracts Section, Institute of Radio Engineers.

"Refrigeration Abstracts," American Society of Refrigerating Engineers.

"Research Abstracts," U. S. National Advisory Committee for Aeronautics.

7 Library Cataloging Systems:

"The Organization of Knowledge in Libraries," by H. E. Bliss, second edition, H. W. Wilson Company, 1939.

"Dewey Decimal Classification," fifteenth edition, Forest Press, 1951.

"Subject Headings," by N. J. Martin, fifth edition, Library of Congress, 1948.

"Subject Headings in Physics," by M. J. Voight, McGraw-Hill Book Company, Inc., New York, N. Y., 1937.

8 Buyers Guides, usually annuals:

"Aeronautical Engineering Catalog," Institute of Aeronautical Sciences.

"ASME Mechanical Catalog," ASME.

"Chemical Engineering Catalog," Reinhold Publishing Company.

"Design News Reference Manual," Rogers Publishing Company.

"Electronics Buyers' Guide," McGraw-Hill Book Company, Inc., New York, N. Y.

"Handbook of Product Design," McGraw-Hill Book Company, Inc., New York, N. Y.

"Instruments Index," Instruments Publishing Company.

"MacRae's Blue Book," MacRae Blue Book Company.

"Nucleonics Buyers' Guide," McGraw-Hill Book Company, Inc., New York, N. Y.

"Product Design File, Sweets' Catalog Service," F. W. Dodge Company.

"Radiation Instrument Catalog," U. S. Atomic Energy Commission.

"Thomas' Register of American Manufacturers," Thomas Publishing Company.

9 AIEE definitions include:

Fundamental and Derived Terms; Rotating Machinery; Transformers, Regulators, Reactors, and Rectifiers; Switching Equipment; Control Equipment; Instruments, Meters, and Meter

Testing; Generation Transmission and Distribution; Transportation, Land and Air; Electromechanical Applications; Electric Welding and Cutting; Feedback-Control Systems; Illuminating Engineering; Electrochemistry and Electrometallurgy; Electrocommunication; Electronics; Radiology; Electrobiology and Electrotherapeutics; Miscellaneous Appliances and Equipment; Automatic Station Control, Supervisory and Telemetering Equipments; Indicating Instruments; Recording Instruments; Industrial Control Apparatus; Electric Railway Control Apparatus; Wire and Cable.

10 ASME definitions include:

Instruments and Apparatus, Power Test Codes; Limits and Fits; Automatic Control; Screw Threads; Shock and Vibration Instrumentation; Fluid Meters; Gears; Surface Quality; High-Pressure Measurement.

11 IRE definitions include:

Antennas, Modulation Systems, and Transmitters; Network Topology; Electroacoustics; Electron Tubes; Magnetrons; Gas-Filled Radiation Counter Tubes; Electronic Computers; Facsimile; Modulation Systems; Radio Aids to Navigation; Piezoelectric Crystals; Radio Receivers; Pulses; Transducers; Color Terms; Television; Wave Propagation; Antennas and Waveguides.

12 Other groups formulating definitions relating to measurement and control:

Acoustical Society of America—Acoustics; Pressure Microphones; Audiometers; Sound Level Meters; Noise.

American Chemical Society—Balances and Weights.

American Medical Association—Atomic Energy and Nuclear Physics.

American Society for Metals—Metallurgy.

American Society for Testing Materials—Color; Acidity; Additions (Cement); Adhesion and Adhesives; Aggregates; Aniline Paint; Ultraviolet Spectrophotometry; Bituminous Waterproofing and Roofing; Calorific Value; Copper-Base Alloys; Cast Iron; Coal and Coke; Ceramic Whiteware; Mortars; Conditioning and Weathering; Panels for Building Construction; Cotton Fibers and Yarns; Cutting Fluids; Density by Pycnometer; Directional Reflectance; Rubber and Synthetic Elastometers; Distillation Range; Drying Oils; Electrical Insulating Oils; Extensometers; Felt; Refractories; Gaseous Fuels; Glass and Glass Products; Industrial Water; Gypsum; Heat-Treatment of Metals; Insulating Materials; Magnetic Testing; Paving Materials; Metallography; Mechanical Testing; Lubricating Oil; Molding Powders; Building Stone; Naval Stores; Traffic Paints; Oxidation Stability; Paints and Varnishes; Paper, Petroleum Products; Plastics; Porcelain Enamel; Portland Cement; Powder Metallurgy; Radiographic Inspection; Rheological Properties; Sand; Shear and Tensile Strength; Shipping Containers; Soaps; Soil Mechanics; Specific Gravity; Specular Gloss; Steel Products; Tall Oils; Thermometers; Textiles; Thermal Conductance and Transmittance; Tile; Water Vapor Permeability; Woods, Veneers and Timber; Wrought Iron; Wove Yarns.

Federal Specification, Government Printing Office—Analytical Balances.

Instrument Society of America.

National Research Council—Nuclear Science and Technology.

National Telemetering Conference—Electrical Telemetering.

Optical Society of America—Photographic Objective Lenses; Radiometry and Photometry; Colorimetry.

Scientific Apparatus Manufacturers of America—Accuracy, Response and Sensitivity of Industrial Instruments.

Society of Applied Spectroscopy.

Society of Automotive Engineers—Lubricants; Bearings; Bolts and Nuts; Brakes and Brake Operation; Crank-Case Oils; Heat-Treating; Hydrodynamic Drives; Serrations and Splines; Pistons and Rings; Ride and Vibration; Screw Threads.

U. S. Department of Defense, Research and Development Board—Guided Missiles.

Discussion

W. B. Heinz.[4] If anyone was ever well qualified to write on a stated subject, the author is qualified to write on terminology. He has participated through the years in the work of the Instruments and Regulators Division Terminology Committee. Through his visits to England and his translations of German papers on automatic control he has had better opportunity than most of us American engineers to incorporate philosophies from other countries in the development of his own. It is his philosophical examination and classification of the terminology problem which outstandingly characterizes his paper.

It is interesting to note the difference he gives between the rights and duties of a lexicographer and those of a terminologist. The lexicographer writes a dictionary as a report on all used words, while the terminologist discriminates among them and creates some, often before usage has come to be. If the paper can stimulate new and wider interest in this important and entertaining work and thus cause others to terminologize before conflicting usages have grown, it may prove of wider benefit than even he himself might hope!

The author's comment on universality of concepts suggests an example which arose early in the years of discussions by the IRD Committee. It was apparent that pneumatic-flow resistance was analogous to electrical resistance and it seemed that definitions should be similar. However, pneumatic circuits are usually nonlinear. That property termed "resistance" was quite naturally recognized as the rate of change of pressure drop with respect to flow, the slope of a line, differing from point to point.

The AIEE Standard Electrical Terms defined resistance as the ratio of potential drop to current. This is correct for a linear element but for a nonlinear one it has quite different significance from the derivative at a specific point. Had AIEE terminologists in those early years given adequate consideration to nonlinear elements their definition would doubtless have been influenced thereby. Their concept might have been more universal.

There is unexpected interest and value in the writing of definitions. Definition gives understanding. Those who compose definitions develop understanding. They also have opportunities to support their pet words and thus to avoid being frustrated by terminology committees who declare that something else should be said! As one who has enjoyed working with the author and his collaborators in developing terminology the writer can honestly urge all those engineers who can write to offer their services and to get into the game.

T. J. Higgins.[5] The writer finds himself in complete agreement with the author's remarks and in complete sympathy with his pleas.

The author's remark, "If the reader is stimulated to lend a hand to the integration of concepts, the author's purpose will have been accomplished," spurred the writer to reread the tentative definitions of terms advanced in "Frequency Response Design Criteria and Standards for Data Presentation," as presented at the Annual Meeting of the ASME in New York, in December, 1953.

He found, confirmative of his recollections, that certain of these definitions deviate widely from already well-established definitions as employed by electrical engineers active in servomechanism and automatic-control theory. Accordingly, he urges that the mentioned definitions—and others stemming from previous, current, and subsequent work by nonelectrical groups—be compared with those now employed in electrical-engineering circles; for as therein the state of the theory of control and regulations is, in general, far in advance of its current state in mechanical-engineering and chemical-engineering circles, it would seem that definitions firmly established by the former could well be taken over wholeheartedly by the latter.

J. B. McMahon.[6] The paper is a well-thought-out presenta-

[4] Proprietor, Heinz Engineering Company, Arlington, Va. Mem. ASME. (Now Chief Engineer, Askania Regulator Company, Chicago, Ill.)

[5] Professor of Electrical Engineering, University of Wisconsin, Madison, Wis.

[6] Manufacturers Representative and Consulting Engineer, Chicago, Ill. Mem. ASME.

tion of a logical approach to the establishment of a terminology in any field. It need not be restricted to that of instrumentation.

There are two points which occur to the writer that might be emphasized more strongly than has been done by the author.

The first is that any definition to be useful must be both inclusive and exclusive. It must include the essentials of the term and also must automatically exclude other uses, which lead to ambiguities. He mentions the term "sensitivity," which has been badly abused by not having been subjected to this treatment.

The second point is that when definitions are phrased to be both inclusive and exclusive, they are apt to sound stiff and stilted. In this respect they become like laws, which are subject to the same qualifications, and on that account sound a trifle queer to lay ears. A long experience in working with the Automatic Control Terminology Committee of ASME has convinced the writer that there is no escape from this situation. If definitions are to be accurate enough for use in purchase specifications, for instance, they probably will be too rigorous for use in casual conversation.

While on that subject, is there any objection to the use of the single word "Instrumentation" in place of the double expression "Measurement and Control?" Is it felt that "Instrumentation" is the art of applying "Measurement and Control," or that "Instrumentation" is an assemblage of equipment? Certainly here is a place where a good definition could effect a simplification which would be helpful.

Author's Closure

Each of the discussers has effectively supported the idea of concept integration, using a specific example from his own experience. Since the paper was presented, ASA has established Sectional Committee C85 on Automatic Control Terminology. Perhaps concepts will prove their value in integrating the deliberations of C85, whose members represent a wide deversity of academic and industrial backgrounds.

Turbulence in a Tilting-Pad Thrust Bearing

By STANLEY ABRAMOVITZ,[1] PHILADELPHIA, PA.

Using water as a lubricant, experiments were made with a tilting-pad thrust bearing. It was found that the friction torque increased abnormally when the rubbing speed exceeded what appeared to be a critical value. This abnormal torque indicated the possible effect of turbulence in the hydrodynamic film between the runner and pads. An analysis was made of experimental data based on the work of G. I. Taylor. These data show the effect on friction torque of the variables of fluid viscosity and sliding speed for both laminar and turbulent film conditions.

Introduction

A SPECIAL thrust-bearing test machine designed and built by The Franklin Institute under subcontract with Westinghouse Electric Corporation was used for this investigation. This machine incorporated an extremely sensitive system for measuring bearing-friction torque. The test bearing was a six-pad, 48-deg, 8-in-OD × 4$^1/_2$-in-ID equalized thrust bearing. The pad pivot was located centrally and the pad surfaces had been lapped to a cylindrical convex-surface configuration. The axis of the cylindrical surface was radial and was of a curvature which represented 0.0003-in. crown.

After the initial indication of excessive torque at a mean bearing velocity of 75 fps, tests were conducted over a range of fluid temperatures and bearing velocities. The water temperature was varied from 75 to 200 F, and the bearing velocity was varied in four steps from approximately 25 to 100 fps. With the torque values obtained, a plot of coefficient of friction versus $\mu U/PB$ was made. The friction deviation from the classic straight-line log-plotted relationship was clearly indicated. It was interesting to find that when the plotted values were grouped for constant temperatures, the result was what appeared to be a family of curves with the fluid temperature (viscosity) as the parameter.

Upon reviewing the literature it was found that G. I. Taylor[2,3] had investigated fluid stability in the annulus between two concentric rotating cylinders. More recently, D. F. Wilcock[4] applied Taylor's method of analysis to the investigation of turbulence in high-speed journal bearings.

In applying Taylor's analysis for a film of constant thickness to a tilting-pad thrust bearing, as we have done, it must be remembered that some error is introduced. The film thickness in a tilting-pad bearing is certainly not constant. The variations in film thickness due to changes in speed and viscosity do produce a variable geometry condition. With regard to the film variation due to pad inclination, the variation is not great; actually for this particular case the slope of the inclined pad is about 0.0003 in. in a distance of 2 in. Other points to consider are surface configuration and interrupted flow between pad segments which do not allow strict comparison to Taylor's work with concentric cylinders. Nevertheless, using Taylor's criterion, some interesting and perhaps significant conclusions may be reached.

Notation

- l = mean length of pad in direction of motion, in.
- b = radial width of pad, in.
- n = number of pads
- h = minimum film thickness, in.
- U = velocity at mean diameter, ips
- P = bearing load, psi
- F = shear force, lb
- μ = absolute viscosity, lb-sec/in.^{2}
- ρ = mass density, lb-sec^2/in.^{4}
- ν = kinematic viscosity, in^2/sec

Derivation of Dimensionless Co-Ordinates

According to Newton's law of viscous flow

Shear stress = (dynamic viscosity) × (rate of deformation)

Therefore, as applied to a tilting-pad thrust bearing

$$F = nlb\left(\frac{\mu U}{h}\right)K_1$$

and

$$\mathrm{Hp} = \frac{K_2\mu U^2}{h} \qquad \text{[1]}$$

where

$$h = \sqrt{0.06\,\frac{\mu U l}{P}}$$

Equation [1] can then be put in the following form

$$\left(\frac{\mathrm{Hp}}{K_2U^2\rho}\right)\left(\frac{Uh}{\nu}\right) = 1 \qquad \text{[2]}$$

Equation [2] is recognized as an equilateral hyperbola. A log-log plot of the two dimensionless relations in parentheses would then yield a straight line with a negative slope of 45 deg.

This method of approach is identical with that of Wilcock[4] as applied to the journal bearing, and stems from the original derivation by Taylor.[2,3]

Description of Thrust-Bearing Test Machine

This machine, Figs. 1 to 5, is designed so that the flat-faced thrust runner is mounted and fixed to the main drive shaft with the bearing face downward. The runner is rested and rotated upon six pivoted thrust shoes mounted upon an equalizer system. The entire bearing is enclosed in a test chamber which is divorced, with the exception of the drive shaft, from driving and auxiliary equipment.

The structure of the testing machine is a four-column stand with two integral decks and a removable subdeck, which is the

---

[1] Project Engineer, The Friction and Lubrication Section, Mechanical Engineering Division, The Franklin Institute Laboratories for Research and Development. Assoc. Mem. ASME.

[2] "Stability of a Viscous Liquid Contained Between Two Rotating Cylinders," by G. I. Taylor, Trans. of the Royal Society of London, England, series A, vol. 223, 1932, p. 289.

[3] "Fluid Friction Between Rotating Cylinders," by G. I. Taylor, Proceedings of the Royal Society, London, England, series A, vol. 157, 1936, p. 546.

[4] "Turbulence in High-Speed Journal Bearings," by D. F. Wilcock, Trans. ASME, vol. 72, 1950, pp. 825–834.

Contributed by the Research Committee on Lubrication under the auspices of the Lubrication Activity of The American Society of Mechanical Engineers, and presented at the First Annual ASME-ASLE Lubrication Conference, Baltimore, Md., October 18–20, 1954.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 3, 1954. Paper No. 54—LUB-7.

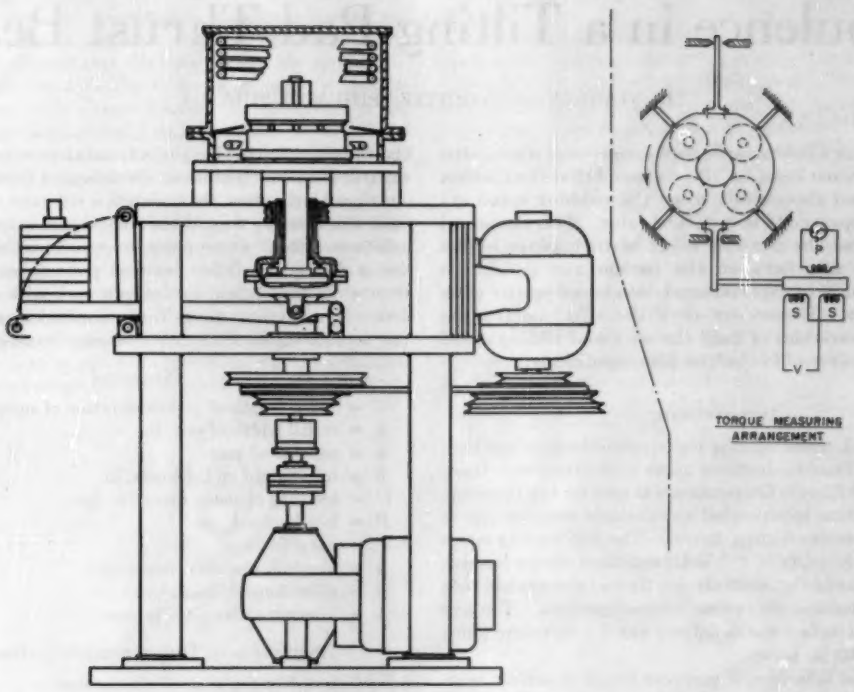


FIG. 1 SCHEMATIC DRAWING OF TEST MACHINE

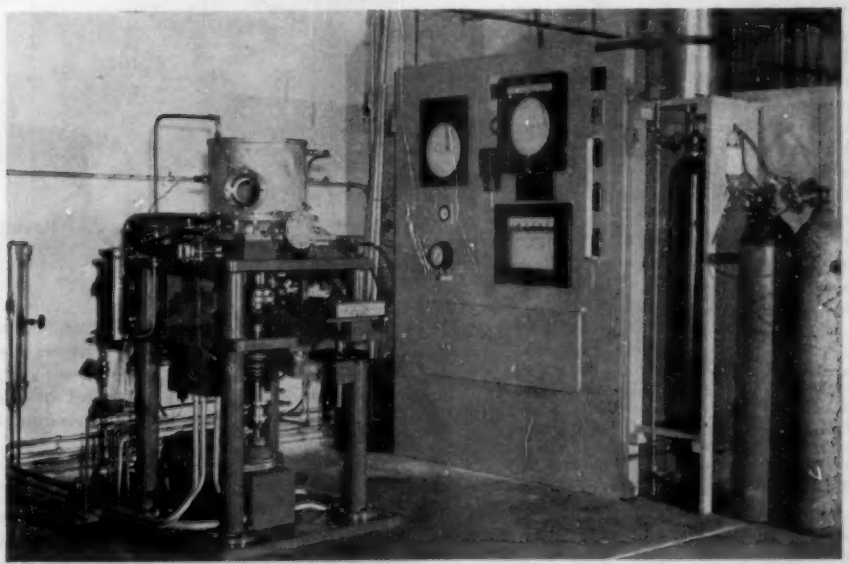


FIG. 2 OVER-ALL PICTURE OF TEST APPARATUS

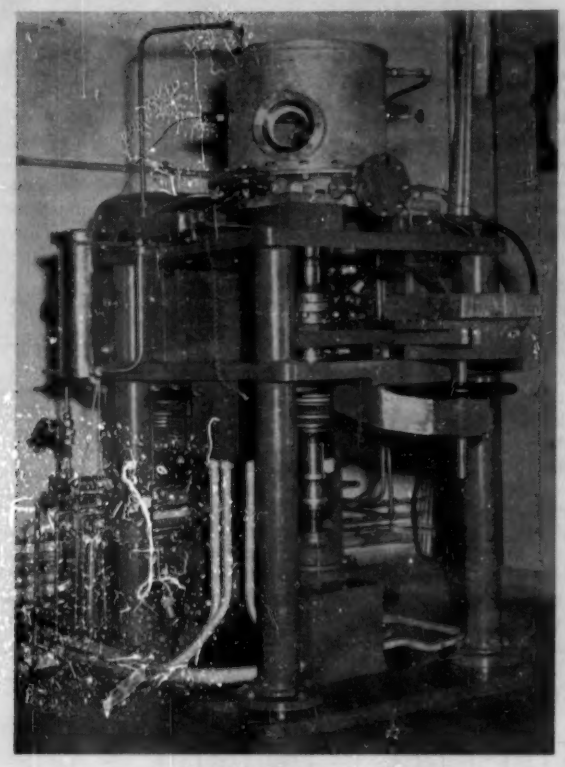


FIG. 3 CLOSE-UP VIEW OF TEST MACHINE

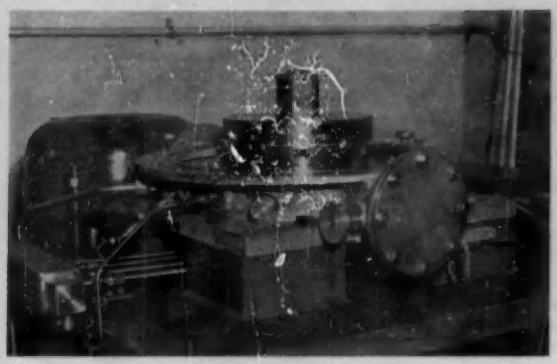


FIG. 4 COMPLETE BEARING MOUNTED IN TEST MACHINE

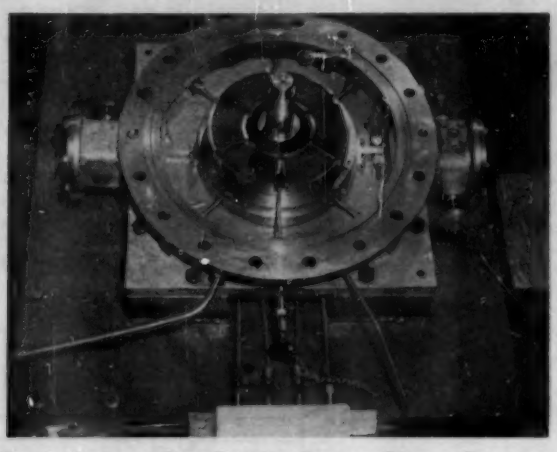


FIG. 5 INSIDE OF TEST CHAMBER SHOWING BEARING SHOES AND TORQUE-MEASURING ARRANGEMENT

base of the test-bearing housing. The centrally located drive shaft is powered by two motors: A low-output-speed gearmotor for high starting torque, and a high-speed motor for running at test speeds which may be varied through step V-belt sheaves. The vertical output shaft of the gearmotor transmits the starting torque through a gear-type coupling to an overrunning clutch. With this arrangement the machine is capable of producing a high starting torque at low speed, after which the high-speed motor may be engaged to drive the shaft at the desired test speed by overrunning the gearmotor.

The thrust load is applied through a load arm linked to an oil-lubricated double-thrust bearing; the bearing runner is keyed and shouldered to the drive shaft. The load-bearing housing is guided between the integral decks to take any small vertical motion desired in the application of load. A cam is provided against the load arm so that the load may be relieved or applied at a given rate. The load bearing is completely divorced from the test housing.

Friction torque is measured by measuring the amount of rotation of an elastically restrained disk which supports the bearing. The disk to which the bearing cage is secured is floated on four water-pressurized circular pads. An external pump and throttling valves maintain a prescribed constant pressure at the pads, by recirculating the water drawn from the test chamber. The disk is held centrally by four radial rods, with spherical ends, equally spaced. Rotation against the elastic restraint is measured by means of a differential transformer whose core is connected to an arm secured to the disk. The elastic restraint of the disk is supplied by a torque bar diametrically opposed to the measurement arm and calibrated so that quantitative torque values can be obtained from the output of the transformer. Stops are provided at the measurement arm to prevent the torque bar from being overstressed when initial high starting torque is developed. A baffle, in the test chamber, is used to prevent the rotating fluid from affecting the torque-measuring disk.

Fluid temperature within the housing is controlled by an electric heater and thermoswitch. Thermocouples located within the housing provide the means for measuring and recording fluid temperatures.

The fluid pressure within the housing is applied and controlled by means of a bladder-type accumulator and a constant-pressure nitrogen supply. The accumulator serves as a replenisher for any leakage from the pressurized system.

Running friction torque, water temperature, and pressure are all continuously recorded. Since it is desired to have the machine operate unattended, safety controls are provided to shut down the machine under conditions of excessive torque, high or low pressure, and high temperature.

TEST RESULTS

Since only four speeds could be attained with the existing test equipment, the fluid temperature was varied at each speed. Thermocouples were located as close to the pads as possible, and a suitable length of time was allowed for equilibrium at the reported temperatures. The bearing was completely immersed in water under pressure of 40 psig. Test data are given in Table 1.

TABLE 1 TEST DATA

Speed, rpm	Fluid temp, deg F	Torque, in-lb	Horsepower	Bearing load, psi
900	87.5	5.2	0.082	23.2
900	110	5.1	0.081	23.2
900	132	4.2	0.067	23.2
900	162	3.7	0.059	23.2
900	198	3.7	0.059	23.2
1750	89	8.5	0.266	21.2
1750	109.5	7.7	0.236	23.2
1755	133	6.7	0.208	23.2
1755	163	6.8	0.211	23.2
1750	198	6.8	0.210	23.2
2600	90	12.5	0.571	23.2
2590	110	12.7	0.58	23.2
2610	133	12.7	0.585	23.2
2610	164	13.8	0.635	23.2
2600	200	14.3	0.655	23.2
3370	110	23.7	1.41	23.2
3410	130	24.8	1.49	23.2
3405	165	24.8	1.49	23.2
3400	200	25.0	1.50	23.2

Discussion and Conclusions

The dimensionless values of Equation [2] are shown plotted in Fig. 6. The experimental points follow the theoretical curve for laminar flow conditions until a critical value of Reynolds number is reached. Above this critical value the horsepower increases and shows what appears to be a transition region, after which the curve tends to approach a horizontal asymptote.

From the experiments by Wilcock[4] a significant but uniform deviation from the laminar flow power characteristics is indicated, whereas in Fig. 6 a region of transition precedes a uniform rise in power. M. C. Shaw,[5] in reviewing Wilcock's paper, states ". . . . . at relatively low speeds the fluid particles have circular paths, as in ordinary journal bearings which operate at normal speeds. As the speed is increased, inertia forces become significant and pairs of ring vortexes are found to appear in the film. As the speed is further increased the orderly vortex motion gives way to turbulence . . . . . one should expect the initial deviation from the laminar friction curve to be due to vortex motion; and a second rather abrupt increase in friction to accompany the beginning of turbulence . . . . ." As indicated by Fig. 6, the power absorbed above the "first critical" Reynolds number increases and then drops sharply. The power then begins to rise gradually and approaches a constant value which appears to be less than the maximum value reached during the initial rise.

The general shape of the plotted curve is similar to the curves obtained by Nikuradse[6] from his investigation of the variation of the resistance coefficient with Reynolds number for artificially roughened pipes. Nikuradse's system of curves is dependent on a pipe-roughness parameter. If a correlation can be made, the effect of bearing surface roughness would be the influencing factor that would define the final shape and magnitude of the Taylor plot above the critical value for film instability.

The bearing materials used for the testing had a surface roughness of approximately 15 microinches rms. The average film thickness was calculated to be from 300 to 900 microinches for the upper and lower values of bearing velocity and fluid viscosity. Assuming an average asperity height of approximately 50 microinches, the magnitude of the ratio of film thickness to roughness then begins to look reasonable in so far as its influence on turbulence. Although no extensive work has been done to determine

[5] Professor of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass.

[6] "Strömungsgesetze in rauhen Rohren," by J. Nikuradse, VDI-Forschungsheft, No. 361, 1933.

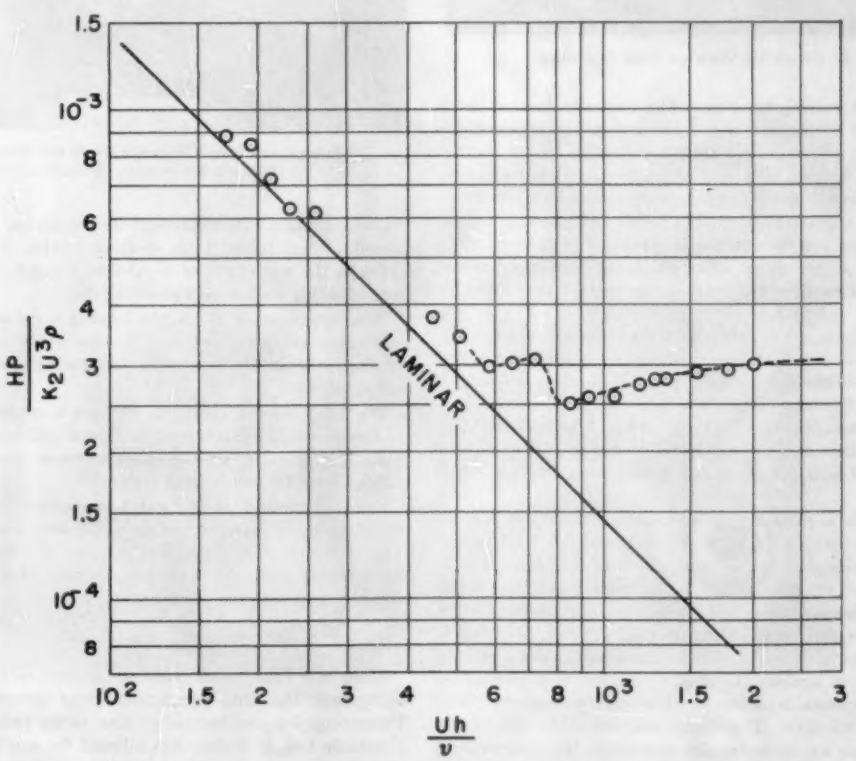


Fig. 6 Taylor Plot for a Tilting-Pad Thrust Bearing Using Water as a Lubricant
($n = 6$; $l = 2.6$ in.; $b = 1.75$ in.)

the influence of surface roughness, some tests with materials having an average roughness of 3 microinches did indicate lower torque values in the turbulent region when compared to the values obtained using the previous material.

With regard to the critical points at which film instability occurs, Fig. 6 shows the start of the transition at a Reynolds number of 580, and the end of the transition at approximately 800. The Reynolds number is based on a mean film thickness, assuming a 2 to 1 ratio of inlet to outlet film thickness. This critical value can be compared with that obtained by D. D. Fuller[7] in his experiments with a journal bearing using liquid Freon 11 as a lubricant. Based on the rubbing velocity at which the torque started to increase abnormally, the calculated Reynolds number was about 970. It is of interest to note that the observed speed at which film instability occurred agreed rather closely with that calculated using Wilcock's equation for the critical point beyond which instability should occur.

Because of the unusual pad characteristics the coefficient used in the film-thickness equation was derived from experimental findings. This was done by first assuming a minimum film thickness based on the height of the surface asperities. Then for a particular unit load the viscosity and minimum bearing velocity were noted when surface contact first occurred. The coefficient was then computed. Actually, this coefficient is not constant for a pad face having a convex profile. A comprehensive mathematical analysis[8] for this general surface configuration disclosed that the load coefficient is a function of the film thickness and pad surface curvature, and will therefore vary with speed, viscosity, and load. However, the variation is quite small and may be neglected for this particular presentation.

Although the deviation of calculated from actual film thickness would affect the numerical values for the critical Reynolds number, the shape of the curve would not be destroyed. In effect, the curve would be translated horizontally.

In conclusion, this investigation has shown an abnormal increase in power above a critical Reynolds number, and has indicated the presence of film instability in a tilting-pad thrust bearing. Actually, further research on this subject is needed. For this particular experiment a more variable speed range would yield added data so that a family of curves with film thickness as the parameter could be obtained. It would be interesting to use different pad and runner materials (glass, for example), and a fluid lubricant other than water but of comparable viscosity. This would be a means of exploring the effects of adhesion, if any. Lastly, the effect of surface roughness should be investigated.

[7] Professor of Mechanical Engineering, Columbia University; Chief, Friction and Lubrication Section, The Franklin Institute.

[8] "Theory for a Slider Bearing With a Convex Pad Surface; Side Flow Neglected," by Stanley Abramovitz, *Journal of The Franklin Institute*, vol. 259, 1955, pp. 221–233.

Investigation of Translatory Fluid Whirl in Vertical Machines

By G. F. BOEKER[1] and B. STERNLICHT[2]

The problem of translatory fluid whirl has been investigated for a rigid shaft running in a vertical position in rigidly supported bearings. This investigation consists of a theoretical bearing analysis and an experimental verification of the theory for two types of bearings.

Introduction

THE literature concerning the vibrations of shafts rotating in journal bearings, where the vibrations are initiated by the fluid-film forces in the bearings, is somewhat confusing. Not only are the reported findings often at variance, but the descriptive terminology is not standardized. We therefore suggest the following definitions:

Whip. This is an instability caused by factors other than the bearing reactions, such as rotor unbalance and the like.

Whirl. This is a motion of the shaft initiated by the hydrodynamic forces within the bearings. The whirl may be of either of two types or a combination of both. It may be an angular whirl which is an angular oscillation of the shaft with respect to the bearing axis or a translatory whirl which is an oscillation of the shaft with its axis remaining essentially parallel to the axis of the bearings.

Whip-Whirl. This is a combination of the mechanical whip and the fluid whirl and is the type of unstable motion frequently observed.

Angular and translatory fluid whirl are illustrated in Fig. 1 of this paper.

Since there are many variables and ramifications associated with the complete problem, such as shaft flexibility, bearing flexibility, gravity loading, misalignment, and the like, it was necessary to limit the investigation to just one phase of the overall problem. The phase that has been investigated is confined to translatory fluid whirl for rigidly mounted bearings operating on a rigid vertical shaft. Therefore, when whirl is mentioned in this paper, translatory fluid whirl will be the subject under consideration. Translatory fluid whirl is actually quite common in situations where vertical machines are used.

Procedure Followed

1 A theoretical analysis predicting the frequency at which translatory fluid whirl will commence for bearings of finite length was developed by modifying a theory proposed by Dr. H. Poritsky for infinite-length bearings.

2 An apparatus was designed and built to measure the threshold whirl frequency of a rigidly supported bearing operating on a stiff vertical shaft, Fig. 2.

3 Two types of bearings were tested in this apparatus. The first was an ungrooved journal bearing while the second was a modified combination-type bearing, Fig. 3.

4 In order to check the theoretically predicted threshold whirl frequency with the experimental one, a determination of the eccentricity locus was also made experimentally. This is a plot of relative position of the shaft center with respect to the bearing center as the operating conditions of the bearing are varied.

5 A comparison was then made between the threshold whirl frequency predicted by the analysis, step 1, and the measured frequency, step 2.

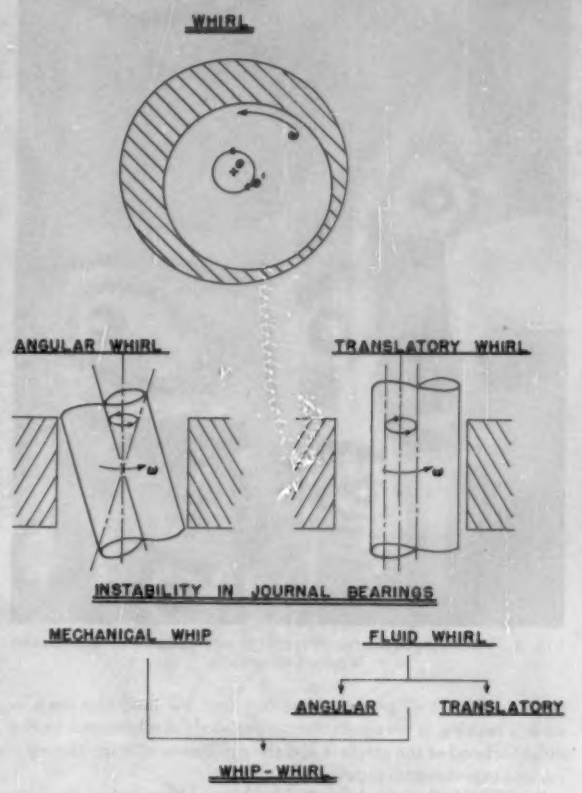


Fig. 1 Illustrative Division of Problem

The comparison between the theoretically predicted threshold whirl frequency and experimentally determined one was made as follows:

For the plain sleeve journal bearing, the experimental eccentricity locus showed that in the neighborhood of the origin the fluid-film force had no component in the direction of the shaft displacement, Fig. 12. The theoretical analysis, step 1, predicted that, under such conditions, the shaft would whirl at all speeds of rotation. This was found to be the case. This experimental result is in agreement with statements of A. C. Hagg,

[1] Department of Mechanical Engineering, Columbia University. At present: The City College of New York, New York, N. Y.

[2] Department of Mechanical Engineering, Columbia University. At present: General Electric Company, Schenectady, N. Y.

Contributed by the Research Committee on Lubrication under the auspices of the Lubrication Activity of The American Society of Mechanical Engineers, and presented at the First Annual ASME-ASLE Lubrication Conference, Baltimore, Md., October 18–20, 1954.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 29, 1954. Paper No. 54—LUB-3

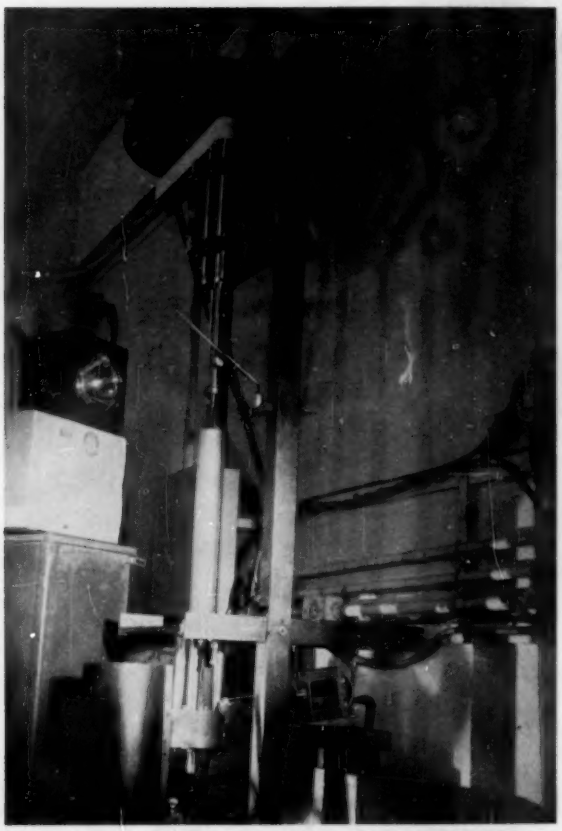


FIG. 2 VERTICAL MACHINE WITH MOVIE CAMERA FOR RECORDING WHIRL FREQUENCY

given in Poritsky's paper.[3] The fact that the fluid-film force in such a bearing is perpendicular to the shaft displacement in the neighborhood of the origin is also the conclusion of many theoretical and experimental papers.

The modified combination type of journal bearing had its characteristics analyzed in several ways. The bearing, consisting of a central part composed of Rayleigh step bearings cut on the shaft and ordinary plain journals at each end, yields to mathematical analysis. The resultant reaction is the vector sum of the reactions due to the Rayleigh step bearing and the plain sleeve bearing. The work of Needs,[4] Dennison,[5] and Wilcock[6] was used to make the calculations. Once the bearing characteristics were known our theoretical analysis could be applied immediately to calculate the threshold whirl frequency.

Fortunately, we were able to get a check on these calculations through the courtesy of the General Electric Company. Using Southwell's relaxation method the Analytical Department also calculated the bearing reactions. Their results and ours were in good agreement.

[3] "Contribution to the Theory of Oil Whip," by H. Poritsky, Trans. ASME, vol. 75, 1953, pp. 1153–1161.

[4] "Effects of Side Leakage in 120 Degree Centrally Supported Journal Bearings," by S. J. Needs, Trans. ASME, vol. 56, 1934, p. 726, Table 2.

[5] "Film-Lubrication Theory and Engine-Bearing Design," by E. S. Dennison, Trans. ASME, vol. 58, 1936, pp. 25–36.

[6] "Oil Flow, Key Factor in Sleeve-Bearing Performance," by D. F. Wilcock and M. Rosenblatt, Trans. ASME, vol. 74, 1952, p. 849.

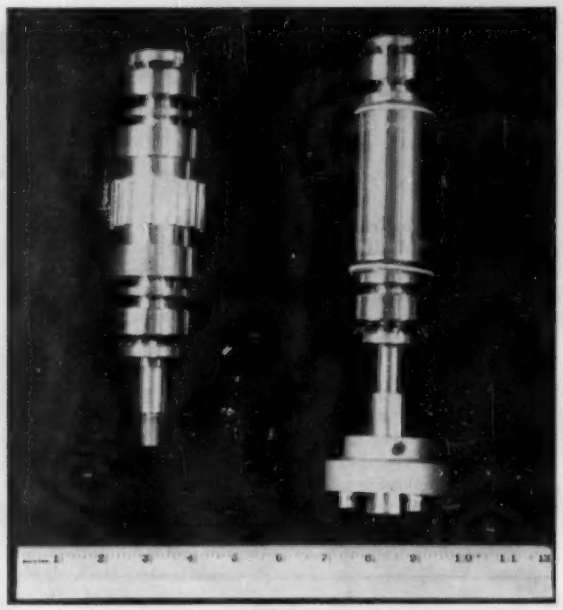


FIG. 3 MODIFIED AND UNGROOVED SHAFT

We also determined the modified characteristics of the bearing experimentally. They checked the theoretically predicted ones quite closely.

THEORY

Determination of Stability Criterion for "Antiwhirl" Bearings. The most recent and probably the best analysis of "oil whip" is that by H. Poritsky.[3] He analyzed the whirl of a full, infinite, plain journal bearing and came to the conclusion that if such a bearing is unloaded and if the fluid-film forces are only those first derived by Harrison,[7] the motion of the rotor is unstable. A finite plain journal, when running in a vertical position, should approximate this condition closely and should thus whirl at all rotational speeds. The experimental work given in this paper seems to verify this conclusion and confirms a similar statement by A. C. Hagg given in the discussion of Poritsky's paper.

Poritsky then showed that stability could be secured mathematically by adding a radial force to those forces given by Harrison's equations. While this is good mathematics, the experimental work shown in this paper seems to indicate that no such radial force exists for an ungrooved full journal bearing in the region of small journal displacement from the bearing axis. It is true that vertical shafts have been run with plain journal bearings without "oil whirl," but probably only because the bearings were not in exact alignment; they were preloaded, or something of the sort, to which Poritsky's analysis does not apply.

Harrison derived the following expressions for the fluid-force components for an infinite, plain journal bearing. In Poritsky's paper they are Equations [2.2] and [2.3]

$$\left.\begin{aligned} F_c &= \frac{12\pi\mu a^2}{\eta_0^2}\,\frac{c}{(2+c^2)(1-c^2)^{1/2}}\left(U-\frac{2aV_\theta}{c\eta_0}\right) \\ &= \frac{12\pi\mu a^2}{\eta_0^2}\,\frac{c}{(2+c^2)(1-c^2)^{1/2}}\,(w-2\Omega) \end{aligned}\right\}. \quad [2.2]$$

[7] "The Hydrodynamic Theory of the Lubrication of a Cylindrical Bearing Under Variable Load and of a Pivot Bearing," by W. J. Harrison, Trans. of the Cambridge Philosophical Society, vol. 22, 1912–1923, pp. 373–388.

$$F_r = \frac{-12\pi\mu a^3}{\eta_0^3} \frac{V_r}{(1-\epsilon^2)^{3/2}} \quad [2.3]$$

Now the section in Poritsky's paper entitled "Effect of Adding Radial Bearing Forces" starts with his Equation [4.1]

$$\left.\begin{aligned} F_x &= -Kw\eta - K_1\xi - 2K\dot{\xi} \\ F_y &= Kw\xi - K_1\eta - 2K\dot{\eta} \end{aligned}\right\} \quad [4.1]$$

for the force due to a small displacement from the origin. They have as their basis Harrison's Equations [2.2] and [2.3] for an infinite plain journal bearing, to which has been added an unknown radial-force component. The Harrison equations are necessary to secure a relation between the coefficients of the terms dependent on the motion of the rotor center and the terms dependent on the position of the rotor center. Hence, even if we were to allow, for the moment, the possibility of a nonexistent radial component in the region about the origin, the result would still only apply to the plain full cylindrical bearings.

The usual method of avoiding oil whip with a vertical unloaded rotor is to use bearings which have some sort of grooving either on the journal or the bearing, to use shoe-type bearings, or to employ other means of a somewhat similar nature. All of these bearings have a radial film-force component in the region about the origin, and are stable at least for all speeds up to some threshold value when run unloaded in a vertical position. Yet we are not justified in applying Poritsky's criterion for oil whip represented in Equation [4.11]

$$Mw^2\left(\frac{1}{k} + \frac{1}{k_1}\right) < 4$$

to them because the basis of his equation, as stated previously, applies only to plain cylindrical bearings and not to these types of "antiwhirl" bearings.

We were aware of the probability of the foregoing facts before this investigation was undertaken. We also were aware of something else, namely, that Poritsky's criterion for bearing whirl checked an experimental finding in a single case where we had accurate data even though the bearing tested was not a plain journal bearing, provided a proper physical interpretation was given to the radial-force component of Poritsky. Hence we were obliged to find some justification of his criterion which does not depend upon the equations of Harrison, which are not applicable to these special bearings.

Justification of Poritsky's Criterion for Bearing Other Than Plain Journal Bearings, Provided They Possess a High Degree of Symmetry With Respect to Their Axis.[5] Bearings such as multiple-pad bearings, shoe bearings having many shoes, bearings having a multiplicity of identical grooves on either journal or bearing can be considered to have sufficient symmetry to make the following argument plausible:

It will be helpful to consider these conditions in vector coordinates. In Fig. 4, O is the bearing center, and O' is the journal center with Cartesian co-ordinates x and y. Let $\hat{\rho}$ be a unit vector from O toward O'. Let $\hat{\theta}$ be a unit vector perpendicular to $\hat{\rho}$ in the direction of increasing θ.

It seems plausible to assume that the film force in the region about the origin is composed of two parts, one due to the motion of the journal center and the other due to the position of the journal center when the shaft is rotating about its own axis, and that the vector equation is of the form

$$\mathbf{F} = -f_1\mathbf{V} + \mathbf{F}_0 \quad [1]$$

where $\mathbf{F}$ is the resultant vector film force on the journal, $\mathbf{V}$ is the vector velocity of the journal center, and $\mathbf{F}_0$ is the film force on the journal resulting from its rotation about its own axis when displaced slightly from the bearing center.

Continuity and symmetry in the vector sense at the origin dictates the form of the first term. The second term can be evaluated experimentally or theoretically and is of the form

$$\mathbf{F}_0 = f_2\hat{\rho} + f_3\hat{\theta} \quad [2]$$

Since

$$\mathbf{V} = \dot{\rho}\rho + \rho\dot{\theta}\hat{\theta} \quad [3]$$

we have

$$\mathbf{F} = [-f_1\dot{\rho} + f_2]\hat{\rho} + [f_3 - f_1\rho\dot{\theta}]\hat{\theta} \quad [4]$$

For small displacements of the center O' with respect to O with a given bearing and lubricant, rotating with constant angular velocity w, we can approximate f_2 by $-K_2\rho$ and f_3 by $K_3\rho$. In general this can be seen by referring to Fig. 5.

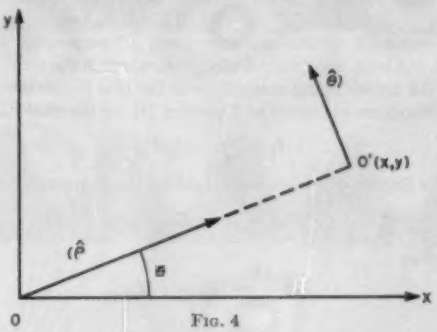


FIG. 4

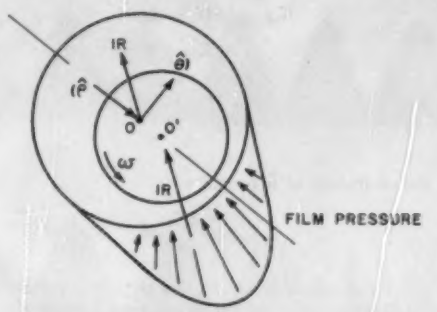


FIG. 5

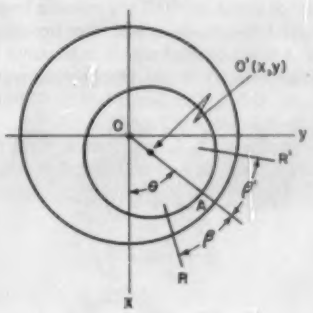


FIG. 6

[5] This also appears in the AEC report NYO-3737 "Preliminary Investigation of Anti Whirl Bearings for Operation in an Acid Solution Containing Abrasive Particles," by C. F. Kayan, G. F. Boeker, D. D. Fuller, R. Corry, B. Sternlicht, and P. Tea (Dec. 15, 1953).

The values of K_2 and K_3 are merely the slopes of the tangent lines to the force components, viz., ρ, at the origin. They can be and were in this investigation found both by known hydrodynamic theory and experimentally.

Term f_1 can be found as follows: Consider Fig. 6.

Let us assume that the spinning shaft is not whirling about the bearing center. It is spinning with angular velocity about its own center, but $\dot{\theta} = 0$. Then the film forces will have a resultant R at some angle β to the line of centers O-O'.

Let us next assume that the journal is whipping with a frequency $\dot{\theta} = w$. Under such circumstances the shaft still has an angular velocity w, so that a point A on the journal always remains outward on the line of centers O-O'. Consider now an observer fixed on the line O-A at O'. He observes the journal to be stationary and the bearing to be rotating with an angular velocity ω opposite in direction to that of the actual shaft rotation.

Since inertia forces and the like have been neglected, and since we assume in the hydrodynamic theory of lubrication that fluid-film forces are due to the relative movement of adjacent surfaces alone, the resultant $\mathbf{R}'$ should be the same in magnitude when the shaft is whipping ($\dot{\theta} = w$) as the resultant for the nonwhirling case R, and the angle $|\beta| = |\beta'|$. It will be seen, however, that the component of this resultant force R' perpendicular to the radius will be in the minus θ-direction, whereas the component of R for the nonwhipping case will be in the plus θ-direction. These components are identified in Equation [4] by the coefficient of $\hat{\theta}$.

$$\mathbf{F}_\theta = (f_3 - f_1\rho\dot{\theta})\hat{\boldsymbol{\theta}} \quad \ldots\ldots [5]$$

For the first case $\dot{\theta} = 0$ (nonwhipping) the component of $\hat{\boldsymbol{\theta}}$ is

$$(F_\theta) = f_3 = K_3\rho = R_\theta \quad \ldots\ldots [6]$$

and when

$$\dot{\theta} = w \quad R_\theta = -R_{\theta'} \quad \ldots\ldots [7]$$

or

$$K_3\rho = -(f_3 - f_1\rho w)$$
$$= -K_3\rho + f_1\rho w$$

hence

$$f_1 = \frac{2K_3}{w} \quad \ldots\ldots [8]$$

Thus the coefficient of $\hat{\boldsymbol{\theta}}$ becomes

$$K_3\rho - \frac{2K_3}{w}\rho\dot{\theta}$$

A check of this equation is afforded by the experimental fact, shown in this research and elsewhere, that a relatively stiff unloaded vertical shaft whirls about the bearing center with a constant frequency of about one-half the running frequency. If $\dot{\theta} = w/2$ the tangential component of the force becomes zero. This is what is required if the journal center is to move in a circle with constant angular velocity about the bearing center as was observed. Actually, this experimental fact could have been used as the basis for the evaluation of f_1.

Having now evaluated f_1, f_2, and f_3, we go back to Equation [1] for $\mathbf{F}$. Using Equation [2] and knowing that

$$\mathbf{V} = \dot{x}i + \dot{y}j$$

$$\hat{\rho} = \frac{x}{\rho}i + \frac{y}{\rho}j$$

$$\hat{\boldsymbol{\theta}} = -\frac{y}{\rho}i + \frac{x}{\rho}j$$

where i and j are unit vectors along the x and y-axis, respectively, $\mathbf{F}$ can be rewritten as

$$\mathbf{F} = \left(-\frac{2K_3}{w}\dot{x} - K_2x - K_1y\right)i + \left(-\frac{2K_3}{w}\dot{y} - K_2y + K_1x\right)j \quad \ldots\ldots [9]$$

Hence

$$\left.\begin{aligned} F_x &= -\frac{2K_3}{w}\dot{x} - K_2x - K_1y \\ F_y &= -\frac{2K_3}{w}\dot{y} - K_2y + K_1x \end{aligned}\right\} \quad \ldots\ldots [10]$$

These equations are exactly equivalent to Poritsky's Equation [4.1] where his K_1 and K are equal to our K_2 and $\frac{K_3}{w}$ upon which the stability criterion

$$Mw^3\left(\frac{1}{k} + \frac{1}{K_2}\right) < 4$$

is based. Here M is the mass of the shaft and $1/k$ is the shaft resilience. Since the stability criterion is solely a direct mathematical consequence of Equations [10], we can adapt Poritsky's criterion for whirl for all such bearings.

It is of interest to note that Poritsky made use of an unpublished theorem[9] in deriving the criterion for whirl from Equations [10]. A derivation not using this theorem and using our Equation [10], applicable to this research, is given in an unclassified report of the U. S. Atomic Energy Commission N.Y.O. 3737.

The experimental work described in this dissertation purposely made use of a relatively stiff shaft so that the factor $1/k$ could be neglected in favor of the $1/K_2$ term. The criterion then becomes

$$w_c = 2\sqrt{\frac{K_2}{M}} \quad \ldots\ldots [11]$$

where w_c is the frequency at which whirl commences.

To summarize the last two sections, we state the following:

Poritsky derived a stability criterion from oil whip for an infinite plain unloaded journal bearing by adding a radial-force component. Unfortunately, no such component exists either theoretically or is found experimentally for a plain journal bearing in the region about the origin. The criterion involves this nonexistent force.

In this paper we have derived a stability criterion for certain types of antiwhirl bearings. Since these bearings are not of the plain-bearing type we could not use as our basis the equations of Harrison as did Poritsky. Hence, if we admitted the possibility of a radial force for a plain journal bearing we still could not use Poritsky's proof. The criterion that we did derive happened to have the same end form as that of Poritsky, but the constants can be evaluated.

Since the mathematics of Poritsky's paper was excellent, we used herein as much of his work as applied to our problem.

Design

As indicated by the analysis, the threshold whirl frequency for a rigid rotor in a vertical machine is a function of the factor K_2, which is the slope of the curve of the radial component of force versus radial displacement, evaluated at the origin.

To verify the analysis, two bearings were designed which had

[9] Poritsky's reference (9)—Internal Report of the General Electric Company, Oct. 18, 1938.

different values of K_3. The two bearings were: (*a*) A plain journal bearing; (*b*) a modified journal bearing. The latter was designed to have a higher value of K_3 so that its threshold whirl frequency would be higher than that for the first bearing. In order to avoid difficulties of fabrication, the shaft was grooved rather than the bearing, Fig. 3.

The design of these bearings was based on the known hydrodynamic theory of lubrication. Calculations were made to determine the load-carrying capacity. This load was further resolved into radial and tangential components with the aid of the eccentricity locus.

Experiment

The frequency and amplitude of whirl for the plain journal bearing were determined from the modulation of a slot whose image was recorded on a photographic film. The lubricant used was air. A schematic representation of this is given in Fig. 7. A photograph of the recording instrument, which was a modified oscillograph, is given in Fig. 8.

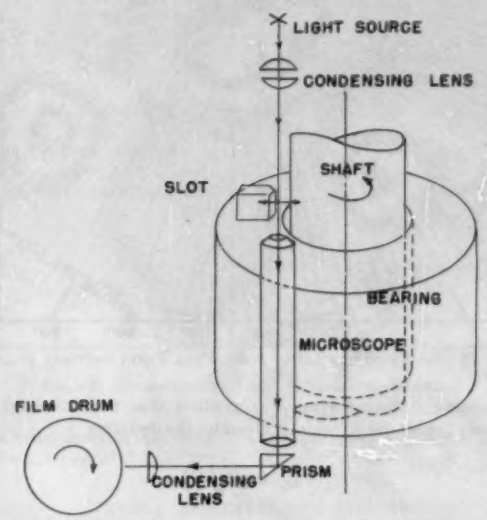


Fig. 7 Schematic Plan for Recording Slot Modulation

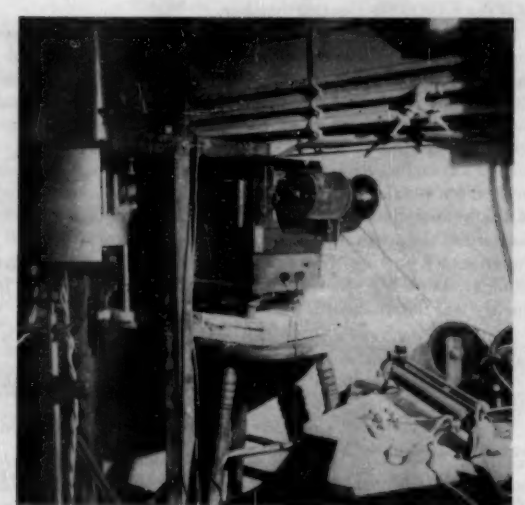


Fig. 8 Modified Oscillograph

Limits between which the threshold whirl frequency lay for the modified journal bearing was determined photographically, Fig. 11. The exact determination was done visually using a strobotac. The lubricant used was distilled water.

Results

The plain journal bearing whirled at all measured speeds when operating in air. The ratio of whirl frequency to shaft rotational frequency was between 0.41 and 0.50. Fig. 9 shows typical modulation data, while Fig. 10 shows a plot of the results. No whirl was detected at speeds lower than 130 rpm on one shaft and 220 rpm on the lighter shaft when the plain journal bearings operated with water as a lubricant.

The modified shaft operated smoothly at a speed of 2620 rpm

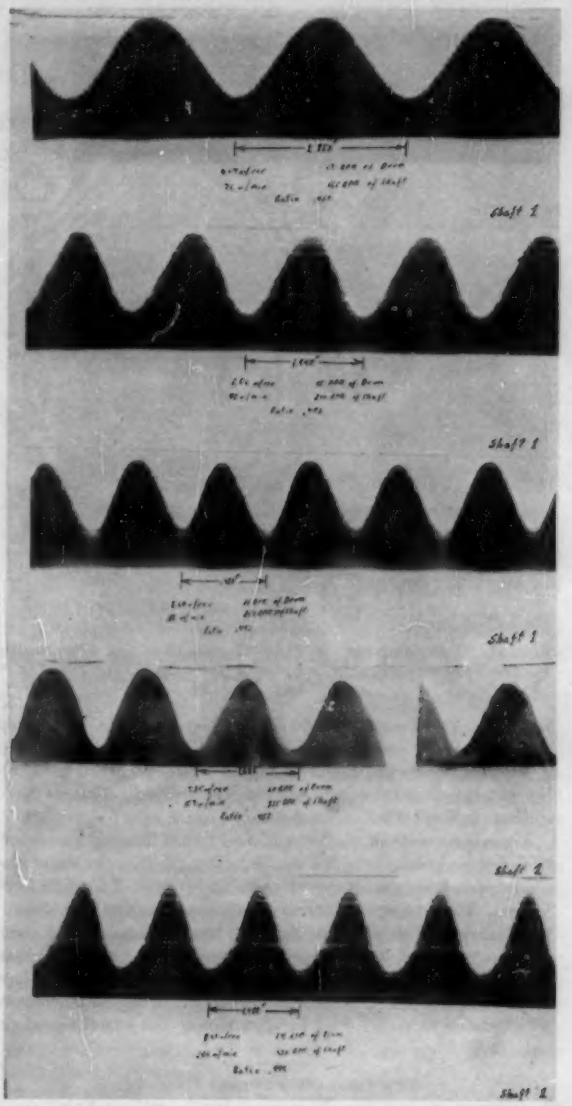


Fig. 9 Oscillographs of Slot Modulation for Determination of Whirl Frequency of Shaft No. 1

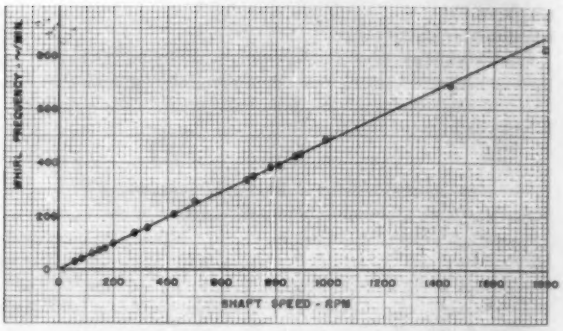


FIG. 10 SHAFT SPEED VERSUS WHIRL FREQUENCY

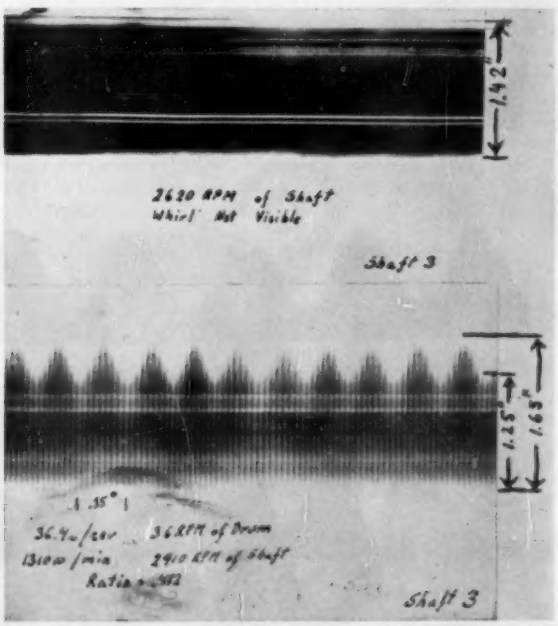


FIG. 11 OSCILLOGRAPHS OF SLOT MODULATION FOR DETERMINATION OF THRESHOLD WHIRL FREQUENCY OF SHAFT NO. 3 WITH WATER USED AS LUBRICANT

but whirled at a speed of 2910 rpm as seen from Fig. 11. Using a strobotac, the initial threshold whirl frequency was found to be 2700 rpm. Once whirl had set in, the shaft speed had to be reduced to about 2400 rpm before the whirl ceased. This seems to indicate a "lag" effect of so ne sort.

The eccentricity locus for a plain journal bearing was determined experimentally. We see from Fig. 12 that for small displacements from the center the radial component of the film force is zero. This indicates that the plain journal bearing should whirl at all operating speeds. On the other hand, the modified journal bearing had a radial and tangential component for small displacements as indicated in Fig. 13. The radial component had a slope of 200 lb/in. obtained graphically and 176 lb/in. as obtained from an algebraic equation used to approximate the experimental results, Fig. 14.

SUMMARY AND FURTHER RESULTS

1 A theoretical analysis of the stability of a very stiff vertical unloaded shaft operating in a rigidly mounted journal bearing

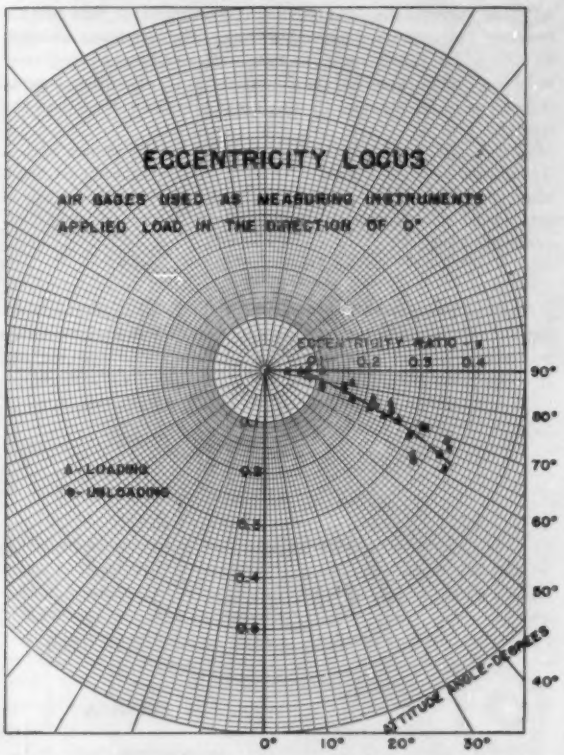


FIG. 12 ECCENTRICITY LOCUS FOR A FULL PLAIN JOURNAL BEARING

possessing sufficient symmetry predicts that the threshold frequency for lateral oil whirl is given by the equation

$$w_c = 2\sqrt{\frac{K_2}{M}}$$

where K_2 is the slope of the tangent to the curve of the radial component of the bearing reaction, viz., journal displacement, and M is the mass of the rotating member.

2 For a plain journal bearing at the origin $K_2 = 0$. This was shown to be the case by experiment. Hence the theory predicts that an unloaded plain journal bearing should whirl at all speeds. Experiments verified this conclusion. A vertically mounted plain journal bearing whirled at all speeds tested. The lowest speed tried was 62 rpm.

3 To get a further verification of the analysis, it was necessary to design a bearing for which the plot of the radial component of the hydrodynamic film force, viz., journal displacement, was not zero at the origin. Known hydrodynamic theory was used to design such a bearing. Three analytical methods were employed to determine the K_2-factor from which the threshold whirl frequency was determined:

(*a*) Based on Need's work[4] the threshold whirl frequency was calculated to be 2890 cpm.

(*b*) Based on Dennison's data[5] the whirl frequency was calculated to be 2260 cpm.

(*c*) Using Southwell's relaxation method, the whirl frequency was calculated to be 2910 cpm.

Methods (*a*) and (*b*) used Wilcock's[6] eccentricity locus for the resolution of the radial and tangential component of force.

The threshold whirl frequency was obtained experimentally for

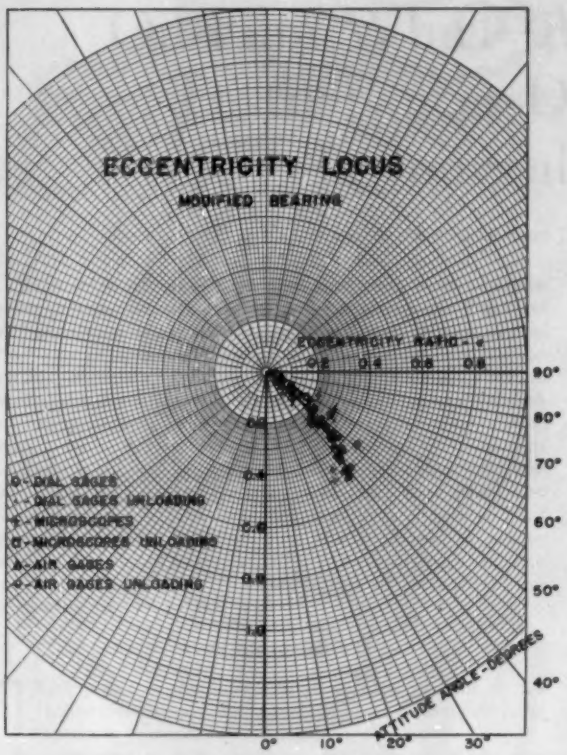


FIG. 13 ECCENTRICITY LOCUS—MODIFIED BEARING

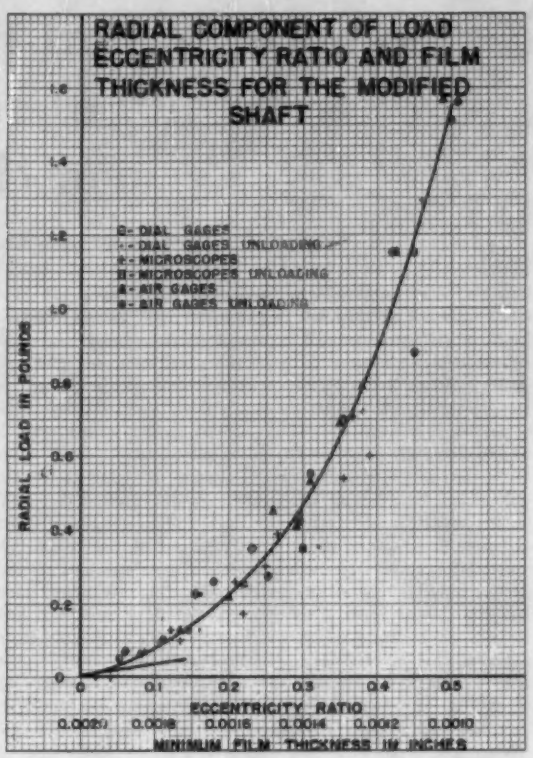


FIG. 14 RADIAL COMPONENT OF LOAD VERSUS ECCENTRICITY RATIO AND FILM THICKNESS FOR MODIFIED SHAFT

this bearing when operating in a vertical position. It was found to be 2700 cpm.

TABLE 1 COMPARISON OF RESULTS

	Threshold whirl frequency			
	Theoretical	Design, cpm	Experimental	Per cent error
Plain journal bearing		0	Whirled at all freq. down to 62 cpm (lowest measurable frequency)	
Modified journal bearing		2260–2910[a]	2700 cpm	16.9–8.1
	$f = \frac{1}{\pi}\sqrt{\frac{K_2}{M}}$	2600[b]		3.7
		2440[c]		9.6

[a] Based on theoretical calculation and references.[4,4,6]
[b] Based on calculations using graphical method for determination of bearing constant K_2.
[c] Based on calculations using algebraic equation for determination of bearing constant K_2.

4 In addition a curve of the radial-film force, viz., journal displacement, for the modified bearing was obtained experimentally. From this curve the slope at the origin was determined, Fig. 14. Based on the slope of the curve, the analysis predicted a threshold whirl frequency using a graphical method for evaluation of the slope of 2600 cpm whereas using an algebraic equation to evaluate the slope, the whirl frequency was calculated to be 2440 cpm.

5 It therefore seems reasonable to assume that this analysis can be used for other types of journal bearings operating under similar conditions provided the bearing form used possesses the requisite symmetry.

ACKNOWLEDGMENT

We gratefully acknowledge the aid and encouragement of Prof. D. D. Fuller and Prof. C. Kayan of the Mechanical Engineering Department of Columbia University, the financial assistance of Higgins Fund, and the General Electric Company, who extended to us many courtesies and some physical aid.

In the main, the experimental work and calculations in this investigation are the work of B. Sternlicht. The theory and some general procedures are the work of G. F. Bocker.

Discussion

J. A. COLE[10] AND C. J. HUGHES.[10] During the course of some film extent experiments on a transparent sleeve bearing (1 × 1 × 0.002 in.) with a single-hole oil entry, we have observed that whirl at frequency near to half shaft speed may occur over a wide speed range but only while the film remains complete. As soon as the film breaks, as a result of increased eccentricity ratio or changed oil-supply conditions, whirl ceases.

A cyclic effect which probably is associated with the construction of the test machine has been noticed. Whirl gradually builds up, then the film breaks down, whereupon whirl ceases and the complete film is formed again, permitting the whole cycle to repeat. If the oil supply is shut off, this state of affairs continues until insufficient oil remains in, or near, the bearing to maintain a complete film.

[10] Lubrication and Wear Division, Department of Scientific and Industrial Research, Mechanical Engineering Research Laboratory, Thorntonhall, Glasgow, Scotland.

Oil-Film Whirl—An Investigation of Disturbances Due to Oil Films in Journal Bearings[1]

By B. L. NEWKIRK[2] and J. F. LEWIS,[3] TROY, N. Y.

With three rotors and five bearings, tests have been run with oil at various viscosities to determine conditions defining a range of stable operation with cylindrical journal bearings at speeds above twice critical speed. It is concluded that short bearings, rather large clearance ratios, and moderate unit bearing loads favor a wide range of such stable operation, up to more than five times critical speed. Slight misalignment resulted in a remarkable increase of the range. After discussion of historical background and consideration of recently reported findings that seem to be at variance with those reported here, it is suggested that two substantially different phenomena should be recognized. Results of a large number of runs under which the whirl was impending (the border line of instability) are presented in two tables.

Introduction

THE object of this investigation is to throw light on the conditions necessary for safe operation of rotating shafts at speeds in excess of twice critical.

Previous work on this subject involves both experiment and theory. The first paper (1),[4] describing the phenomena, reported experimental investigations, but Harrison (2) had already pointed out that the hydrodynamic equations for a full bearing with load, but without end leakage, admit solutions indicating a whirling of the journal in a closed orbit within its clearance. This has since been confirmed (3), and it appears that the indicated frequency of such a whirl is slightly less than half running speed.

Prof. David Robertson (4) remarked that the centrifugal force due to the mass of the journal, whirling in such an orbit, would cause the journal to spiral outward, indicating instability of journal bearings at all speeds.

Poritsky (5), noting that journal bearings are usually stable at speeds below twice critical and viewing other discrepancies between theory and observed behavior, assumed the existence of a radial inward force exerted by the oil on the journal. When terms representing such a force were inserted into his equations they indicated (for small eccentricity) stability at speeds below twice critical and instability at higher speeds.

With light loads, or no load, as in guide bearings of vertical shafts, the phenomena described by Harrison, Swift (3), and Robertson occur. The frequency of the whirl is sometimes less than half-running-speed frequency. The unit bearing loads of our apparatus are not extremely low. However we have observed the half-running-speed-frequency whirl. Half-running-speed-frequency shaft vibration has been observed, with running speed well over twice critical speed. These disturbances are easily damped out. They do not build up resonant vibration of the shaft except when the running speed is twice critical speed.

There is a more severe disturbance that builds up in the natural frequency of the shaft for transverse vibration. This disturbance has not built up in our tests unless shaft speed is equal to, or more than, twice critical speed. With some rotor-bearing combinations and low oil viscosities the lowest speed at which this form of instability sets in is more than five times critical speed. In other cases it is little more than twice critical speed. We have called the lowest speed at which the disturbance builds up spontaneously during the normal quiet running of a well-balanced rotor the "whirl-impending" speed.

This disturbance builds up spontaneously at all speeds within the range of our apparatus above the whirl-impending speed. The frequency of the vibrations, from the start of the build-up, is that of the natural frequency of the shaft for transverse vibration in the running condition. (The lowest natural frequency of the shaft for transverse vibration varies slightly with length and alignment of bearing and elasticity of the oil film when the shaft is running.) The stimulus seems to increase with increasing speed and there seems to be no stable range at higher speed.

The phenomenon described in the last two paragraphs is substantially different from the half-running-speed-frequency disturbance described in the preceding paragraph, although both are due to oil-film action.

Speeds on Boundary Line of Stability

Most of our attention has been devoted to determining the speeds on the boundary line of stability (whirl-impending speeds) for this phenomenon, and the running positions of the journals, and other conditions, at such speeds. The conditions include bearing diameter, clearance, load, length-diameter ratio, and oil viscosity.

It was found that in some cases the whirl-impending speed did not indicate a condition of great stability, since a jar would cause the severe disturbance to build up at a lower speed. A wooden block was set up, supported by wires to swing as a pendulum and give a transverse bump of known momentum to the rotor. The speed at which the rotor would recover and settle down to smooth operation after this bump was called the "bumping speed" and these speeds give better criteria for safe operation than the whirl-impending speeds, which may give better indications for theoretical study of the instability.

We have three rotors and five bearings. The journal diameters are approximately equal, so that all five bearings may be used with each rotor. There are three bumper weights adjusted to the three rotors according to a dimensionless expression

---

[1] This investigation was made under contract with the Air Force and a complete report will appear as WADC-TR54-188.

[2] Consulting Engineer and Professor Emeritus, Rensselaer Polytechnic Institute. Fellow ASME.

[3] Rensselaer Polytechnic Institute.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Lubrication under the auspices of the Lubrication Activity of The American Society of Mechanical Engineers and presented at the first annual ASME-ASLE Lubrication Conference, Baltimore, Md., October 18–20, 1954.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, July 28, 1954. Paper No. 54—LUB-4.

$$\frac{mv}{M\,\omega_c\,L}$$

correlating the momentum of the bump with the mass, stiffness, and span length of the rotor. Here mv represents the momentum of the bump, including recoil, M and L are the mass of the rotor and the span length between bearings, and ω_c is the first critical speed of the rotor.

When the rotor runs with small eccentricity ratio there is sometimes a considerable difference between the "quiet" whirl-impending speed and the bumping speed, but with large eccentricity ratios the two speeds are identical. No separate "bumping" runs were necessary for rotor No. 3.

The Apparatus

The apparatus is mounted on a concrete pier 27 in. wide at the bottom and 27 in. high, poured on the concrete ground floor of the Vibrations Laboratory. Bearing standards are bolted to a 1-in. steel plate which is grouted and bolted to the pier. Attention has been given throughout to rigidity and avoidance of damping. The rotors are balanced accurately.

Fig. 1 shows the test apparatus with the enlarging camera in position and the high-speed driving motor at the farther end. Fig. 2 shows a close-up of the test-bearing standard, the microscope for visual observation, and the adjustable pointer. The other end of the rotor is supported by a self-aligning precision ball bearing.

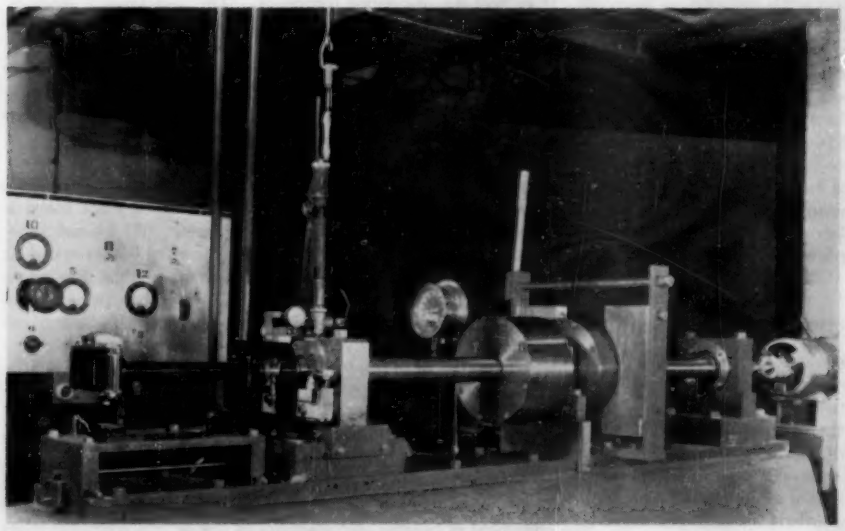


Fig. 1 Test Apparatus

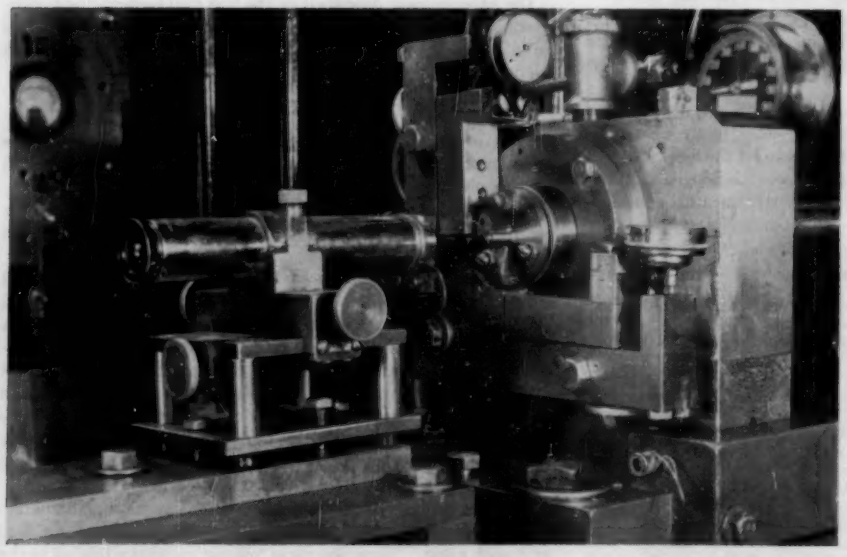


Fig 2 Close-Up of Test-Bearing Standard Microscope and Adjustable Pointer

The rotor is driven by a motor rated at 4 hp at 8000 rpm. The drive is direct, through a Lord rubber coupling. By varying field and armature voltages we are able to drive the rotors, ranging in weight from 192 to 590 lb, at speeds from 1400 to over 8500 rpm with good control at all speeds.

The bearings are cylindrical except for an oilhole and axial oil-distributing groove at the top. Oil is supplied under a gravity head of 11 in. The upper $4^1/_2$ in. of the bearing standard, consisting of two bars bolted together, constitutes a block into which the bearing is pressed before boring. There are five such blocks with bearings of different lengths and clearances.

The shaft end projects through the bearing a distance of $1^1/_4$ in. Attached to the bearing standard is an assembly of two dial gages with scale-multiplying arms to read horizontal and vertical displacements of the shaft.

In the end of the shaft is mounted a pointer with a steel ball that acts as a convex mirror. A Point-O-Lite lamp is set up about 10 ft from the pointer and on a line about 45 deg from the shaft axis. The pointer is adjustable to bring the steel-ball mirror into line with the shaft axis, so that when the rotor is running steadily the image of the light appears as a point. When the journal vibrates or whirls the image appears as a line or loop.

Oil is circulated through a cooler that may be used as a heater also.

A brake is used to quiet the rotor when violent whirling

sets in, and brings the rotor to rest quickly when fiducial points are to be photographed, as will be explained.

There is a thermocouple in the entering oil; four more are in the bearing, and three more are on the bearing standard.

The camera is an Exacta with the lens replaced with a Raptar 16-mm movie camera lens of 1/2-in. focal length, on a 12-in-long barrel, giving a magnification of 26×.

The journals are 2 in. diam, unit loads from 24 to 146 psi, length-diameter ratios from 1/2 to 1, clearance ratios, 2 and 4 parts per thousand, and oil viscosities from less than 1 microreyn to over 7.

The Observations

After a good deal of preliminary study a program was laid out for determining conditions on the border line of instability. As stated previously there are three rotors, all of the same length, but varying in weight and in stiffness. There are five bearings, varying in length and clearance. Runs are made with oil entering at four different oil-inlet temperatures for each rotor-bearing combination.

The procedure is as follows: A rotor is set up with one of the bearings, and the bearing is aligned with care so that the journal lies along the bottom of the bearing, with the pressure distributed as evenly as possible along the line of contact. This involves a tipping of the bearing to match the slope of the rotor. Accuracy of this adjustment is essential.

The pointer is adjusted so that the ball runs true, and the reflected image of the light source appears in the microscope as a small dot with the rotor turning over at a moderate speed. The ball is wiped off with a nylon cloth and the camera is put in place and focused.

For the first runs the cooling water is turned on, the oil pump is started, and the rotor run up to a speed where a disturbance starts. Speed is then reduced slightly so that the disturbance drops out, and operation continued for 1/2 hr or so to establish temperature gradients in the bearing and the bearing standard.

Speed is then increased in steps of 25, 50, or more rpm until a speed is determined at which traces of vibration at the natural frequency of the shaft for transverse vibration appear but do not build up. This is called the whirl-impending speed for quiet operation. These traces of vibration are observed by holding a light forked stick on the shaft and holding a Frahm tachometer on top of the stick. A slight vibration of the reed corresponding to the shaft frequency indicates the start of the unstable motion. In the cases where the eccentricity ratio is large, as it is with the heaviest rotor, there is often a speed range within which it is difficult to decide what the whirl-impending speed is. It is usually possible to determine within 50 rpm a speed above which the vibration builds up to a severe disturbance and below which it remains negligible.

When the whirl-impending speed has been determined and established, any trace of vibration is damped out with the shaft stick and an exposure is made. The thermocouple readings are then made.

The rotor is then brought to rest quickly by means of the brake, and four more exposures are made with the journal lying in the bottom of the bearing, the rotor being turned 90 deg between exposures.

A steel strap is then applied close to the bearing, with elastic support, to take the weight off the bearing so that the journal can be moved around within the bearing with light hand pressure. Fiducial points are then recorded by making exposures with the journal at the bottom, top, and two sides of its clearance travel. Four exposures are made in each position, the rotor being turned 90 deg between exposures.

This program of exposures for fiducial points, including those made when the rotor comes to rest, with the rotor lying in the bottom of the bearing, is completed within about 150 sec from the time power is shut off.

Three runs as just described are made for each temperature. Values of eccentricity ratio and attitude angle are determined from the three photographs and averaged. For each rotor-bearing combination, runs are made at four oil-inlet temperatures, viz., cold oil (depending on cooling-water temperature), 100 F, 130 F, and 160 F.

The exposures are made on 35-mm film with an enlargement of 26×. Before the film strip is cut up, crease lines are drawn on the film side across the picture fields, normal to the film edges. These lines are the basis for determining the angles φ.

The film strip is cut up and the negatives are mounted for projection in Kodak "readymounts."

The pictures are then projected from a fixed distance onto an $8\frac{1}{2} \times 11$ sheet of white paper, and centers of the images, and the creased line, marked down with a pencil. The total magnification is 700×.

Certain corrections are necessary because of the fact that the lifting force is exerted by the strap at a place inside the bearing, thus shortening the span of the rotor and changing its slope. At the higher temperatures (130 and 160 F) small corrections due to cooling down of the bearing standard are required. Details are omitted but they will be given in full in the technical report.

The bumping runs were made as described, except for the running speeds which were the highest speeds at which the rotor would quiet down after the bump. In most cases the speeds at which the rotor would recover from the bump were so close to the whirl-impending speed that separate bumping runs were unnecessary.

Table 1 presents the results of quiet runs with rotor No. 1, the lightest of the three. Table 2 shows results of bumping runs with the same rotor. Table 3 shows results for rotor No. 3. At this writing runs with rotor No. 2 remain to be made.

The ratio of whirl-impending speeds to critical speeds ranges from 2 to 5.9. For the bumping runs the ratio is from 2 to 5.

Discussion

Several features of the disturbance phenomena are apparent.

With the lighter rotor the highest ratios of whirl-impending speed to critical speed (ω/ω_c) were reached with the shortest bearing and the largest clearance. Unit load was 48 psi and the eccentricity ratio was approximately 0.6. With somewhat warmer oil, whirl-impending speeds were attained that were well over five times the critical speed.

With the heavier (No. 3) rotor, higher ratios of whirl-impending to critical speed were found for conditions giving less than the maximum eccentricity ratios. Highest whirl-impending speeds (up to nearly six times critical speed) were attained with an L/D ratio of 3/4 and a clearance of four parts per thousand. The load was 98 psi and the eccentricity ratio 0.85.

In all cases with No. 3 rotor, warming the oil increased the range of stable operation.

In several cases it was noted that slight misalignment of the bearing, such that the journal tended to bear more heavily on the inner end of the bearing, resulted in much higher whirl-impending values.

In some cases, notably with the No. 1 (light) rotor, the $1\frac{1}{2}$-in-long bearing and the four-parts-per-thousand clearance ratio, the half-running-speed-frequency whirl was frequently in evidence. It was easily damped out, and it came and went erratically. It was moderately severe, but it did not build up in violence as the more usual form of instability did. Its frequency was exactly half the running frequency as nearly as could be determined with Frahm tachometers. The critical-speed-frequency disturbance

TABLE 1 ROTOR NO. 1—QUIET RUNS

Log Sheet No.	Bearing	psi	μ c.s.	$\mu \times 10^6$ Reyns	e/C	φ deg	ω rpm	ω_c Dist Freq	$\frac{\omega}{\omega_c}$
48 D-F	7-1-2	48.2	26.0	3.45	.57	74	3375	1600	2.1
48 G-I	"	"	16.0	2.12	.62	70	3600	1610	2.2
48 J-L	"	"	10.7	1.42	.68	67	4100	1610	2.5
48 M-O	"	"	6.4	.85	.67	63	5900	1620	3.6
49 A-C	9-1-4	48.2	39.5	5.24	.58	59	4600	1580	2.9
49 D-F	"	"	27.0	3.58	.60	58	5900	1590	3.7
49 G-I	"	"	22.5	2.99	.62	58	7000	1600	4.4
51 D-F	"	"	17.2	2.28	.64	56	7900		4.9
52 A-C	8-1½-2	32.1	19.5	2.58	.31	83	3350	1620	2.1
52 D-F	"	"	14.5	1.92	.37	81	3350	1620	2.1
52 G-I	"	"	10.5	1.38	.43	77	3475	1610	2.2
52 J-L	"	"	7.4	.98	.57	69	3350	1620	2.1
54 A-C	10-1½-4	32.1	34.0	4.51	.42	69	3350	1580	2.1
54 D-F	"	"	19.5	2.59	.56	66	3325	1590	2.1
54 G-I	"	"	11.0	1.46	.68	56	3550	1595	2.2
54 J-L	"	"	6.3	.84	.71	56	5400	1595	3.4
46 A-C	12-2-4	24.1	33.0	4.38	.31	78	3350	1600	2.1
46 D-F	"	"	19.4	2.57	.44	76	3275	1580	2.1
46 G-I	"	"	12.0	1.59	.60	66	3225	1580	2.0
46 J-L	"	"	6.5	.86	.79	57	3125	1580	2.0

TABLE 2 ROTOR NO. 1—BUMPING RUNS

Log Sheet No.	Bearing	psi	μ c.s.	$\mu \times 10^6$ Reyns	e/C	φ deg	ω rpm	ω_c Dist Freq	$\frac{\omega}{\omega_c}$
48 D-F	7-1-2	48.2						1600	2.1
48 G-I	"	"						1610	2.2
48 J-L	"	"						1610	2.5
48 M-O	"	"					5100	1620	3.1
51 A-C	9-1-4	48.2	39.5	5.24	.61	59	3800	1580	2.4
50 D-F	"	"	31.0	4.12	.66	55	4100	1580	2.6
50 G-I	"	"	26.0	3.46	.66	52	4500	1580	2.8
50 J-L	"	"	18.5	2.46	.69	49	5050	1580	3.2
50 M-O	"	"	6.8	.90	.75	42	8100	1600	5.0
53 A-C	8-1½-2	32.1					3325	1620	2.0
53 D-F	"	"					3325	1620	2.0
53 G-I	"	"	11.3	1.50	.42	80	3275	1620	2.0
53 J-L	"	"					3300	1620	2.0
55 A-C	10-1½-4	32.1					3350	1580	2.1
55 D-F	"	"					3275	1580	2.1
55 G-I	"	"					3550	1595	2.2
55 J-L	"	"	6.3	.84	.75	51	4250	1595	2.7
46 A-C	12-2-4	24.1							2.1
46 D-F	"	"							2.1
46 G-I	"	"							2.0
46 J-L	"	"							2.0

was observed simultaneously with the half-running-speed-frequency disturbance.

In general, the phenomena observed in this investigation differ from the half-running-speed-frequency whirls reported in the early papers and more recently by Pinkus (6) in the following respects:

(*a*) These phenomena are not observed at running speeds below twice the first rotor critical speed.

(*b*) The frequency of these disturbances is always equal to the natural frequency of the rotor for the running condition.

(*c*) When the disturbance starts at any rotor speed it continues with increasing vigor for greater rotor speeds, and it does not quiet down at higher speeds within our range of operation.

(*d*) These disturbances build up on cold oil at lower speeds than they do on hot oil.

(*e*) For any given rotor-bearing combination the eccentricities increase as the oil viscosities decrease and whirl-impending speeds increase.

There are some exceptions to items (*d*) and (*e*) in the case of the lightest rotor and the longer bearings. In such cases small eccentricities predominate and there is very little, if any, range of stable operation above twice critical speed.

It appears that there are two substantially different phenomena to which the terms "oil whip" and "oil-film whirl" have been applied indiscriminately. One of these is a tendency of a journal to whirl at any speed, as discussed by Harrison (2), Swift (3), Robertson (4), Shaw and Macks (7), Burwell (8), and others. The other is an instability feature of the oil film in combination with elastic properties of the rotor that occurs under certain conditions of load, speed, oil viscosity, etc., resulting in the building up of a self-excited resonant vibration. The former phenomenon may set in at running speeds below twice critical speed and pass over into the latter as speed increases through twice critical. These phenomena were observed but not so clearly discerned or described in previous papers (14, 15, 16). References to other papers are given in the Bibliography (9, 11, 12, 13).

Conclusion

In the course of the years since the first published notice of the oil-film whirl, it has been found that various bearing modifications favor stable operation at speeds above twice critical. Among these are the "more stable bearing" described in references (15) and (16) and the "elliptical bearing" so-called. The latter bearing is shaped by placing shims between the two halves, boring the bearing to a large clearance, and then assembling the bearing in the machine without the shims. The more stable bearing is designed so that oil pressure builds up in the top half, exerting a downward pressure on the journal. It has been found also that the disturbance sometimes occurs in an operating machine when the oil is too cold and disappears when the oil is warmed up.

These cases have one factor in common, namely, an increase in eccentricity ratio, accompanied by increased stability. Hummel (10) studied the half-bearing and concluded that a mild instability develops when the eccentricity ratio becomes less than 0.7. These indications suggested that there might be some limiting value of eccentricity ratio beyond which operation would be stable. This seems not to be the case, since we have found instability with eccentricity ratios up to 0.96. However, plots of eccentricity ratio versus whirl-impending speed for any rotor-

TABLE 3 ROTOR NO. 3

Log Sheet No.	Bearing	psi	μ c.s.	μx10^6 Reyns	e/C	φ deg	ω rpm	ω_c Dist Freq	$\frac{\omega}{\omega_c}$
38 A-C	7-1-2	148	20.8	2.76	.80	42	2600		2.5
38 D-F	"	"	13.0	1.72	.82	34	3250		3.1
38 G-I	"	"	8.0	1.06	.85	29	3500		3.4
38 J-L	"	"	5.4	.72	.89	26	4200		4.0
41 A-C	9-1-4	148	55.0	7.30	.82	36	2250		2.2
41 D-F	"	"	21.0	2.80	.90	24	2550		2.5
41 G-I	"	"	12.7	1.68	.93	20	2700		2.6
41 J-L	"	"	6.9	.92	.95	17	3175		3.1
39 A-C	8-1½-2	98	30.5	4.05	.59	65	2300		2.2
39 D-F	"	"	19.3	2.56	.70	56	2450		2.4
39 G-I	"	"	12.4	1.64	.79	46	2600		2.5
39 J-L	"	"	7.6	1.00	.80	42	2950		2.9
43 A-C	10-1½-4	98	55.2	7.30	.69	55	2400	1030	2.3
43 D-F	"	"	18.0	2.40	.82	40	2950	1040	2.8
43 G-I	"	"	10.0	1.33	.85	34	3950	1050	3.8
43 J-L	"	"	6.0	.80	.85	34	6200		5.9
45 A-C	12-2-4	74	50.6	6.70	.66	64	2450	1020	2.4
45 D-F	"	"	18.2	2.40	.89	50	2750	1030	2.7
45 G-I	"	"	10.5	1.40	.94	44	3125	1040	3.0
45 J-L	"	"	6.3	.84	.97	42	3900	1050	3.7

bearing combination and for successively increasing oil temperatures, frequently show asymptotic approaches to definite values of eccentricity ratio. These values are, however, different for different rotor-bearing combinations.

The eccentricity ratio is an important criterion of stability, but the L/D ratio, unit load, and clearance ratio, are significant also, as a glance at the tables will show. The great effect of small changes of alignment is also noteworthy. Moderate unit loading seems desirable for, although stability may be attained with high unit loading, it involves operation on thin films.

At this writing the experimental program is not complete. Runs are to be made with a No. 2 rotor, having weight and critical-speed values between the other two. The tabulated data will be presented with further details of apparatus and procedure, and discussion in Technical Report WADC-TR54-188.

Tables

In the tables, column 1 refers to the log sheet, for use in case of further investigation of any item at some future time. Column 2 describes the bearing; 7-1-2 means that it is the seventh bearing, having a length of 1 in. and a clearance of two parts per thousand. Column 3 gives the load per square inch of projected bearing area. Columns 4 and 5 give the viscosity of the lubricant μ in centistokes and microreyns, respectively. Columns 6 and 7 give the eccentricity ratio and attitude angle of the journal. Column 8 gives the running speed at which the photograph was made. Column 9 gives the disturbance frequency observed with the tachometer for the test conditions, and column 10 gives the ratio of the running speed to the disturbance frequency.

Figs. 3 and 4 show plots of ω/ω_c versus μ for the quiet runs.

It appears that in most cases operation is not stable at speeds much above twice critical speed (values of $\omega/\omega_c > 2$) except with

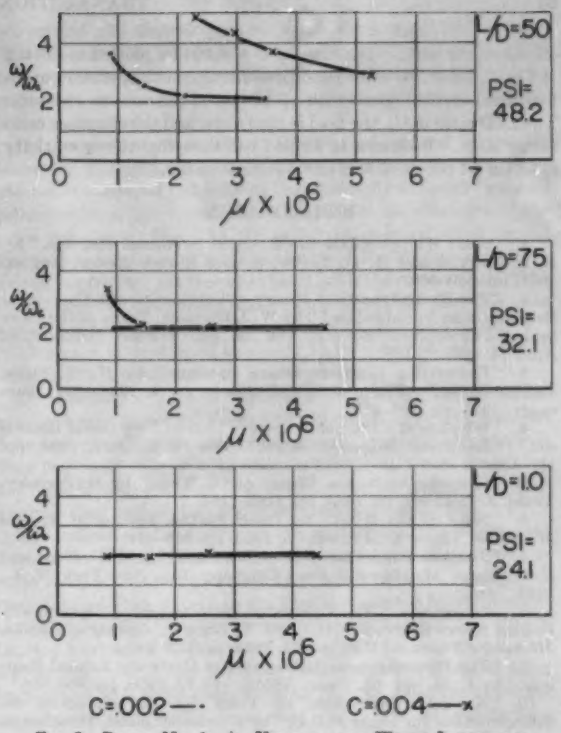


Fig. 3 Rotor No. 1 ω/ω_c Versus μ for Whirl Impending

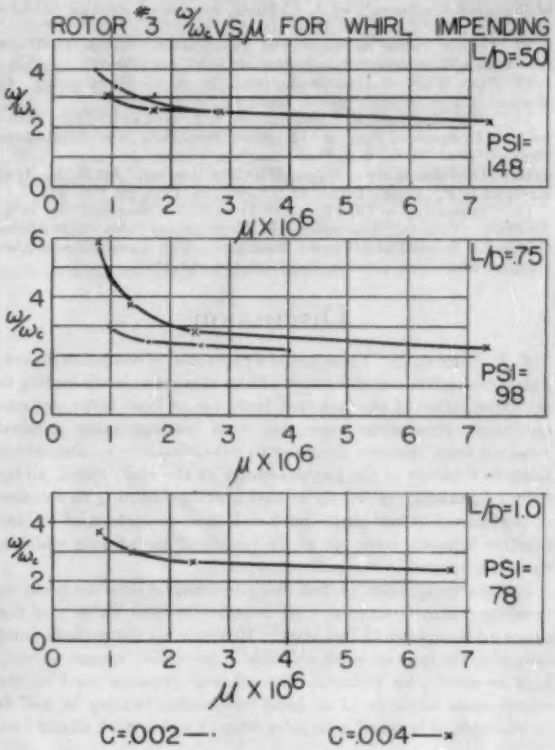


Fig. 4 Rotor No. 3 ω/ω_c Versus μ for Whirl Impending

oil of low viscosity. The runs for $C = 0.004$ in. plotted at the top of Fig. 3, however, show conditions favorable to operation up to five times critical speed with oil of over 2-microreyns viscosity. The L/D ratio is $^1/_2$, the load is moderate, and the clearance ratio rather large. Reference to Table 1 will show that the eccentricity ratio for all points of this curve is close to 0.6.

BIBLIOGRAPHY

1 "Shaft Whipping Due to Oil Action in Journal Bearings," by B. L. Newkirk and H. D. Taylor, *General Electric Review*, August, 1925, pp. 559–568.

2 "The Hydrodynamical Theory of Lubrication of a Cylindrical Bearing Under Variable Load," by W. J. Harrison, Trans. of the Cambridge Philosophical Society, vol. 22, pp. 373-388 (1912–1923), April 24, 1919.

3 "Fluctuating Loads in Sleeve Bearings," by H. W. Swift, *Journal of the Institution of Civil Engineers*, no. 4, 1936–1937; February, 1937, pp. 161–195.

4 "Whirling of a Journal in a Sleeve Bearing," by David Robertson, *Philosophical Magazine*, series 7, vol. 15, January, 1933, pp. 113–130.

5 "Contribution to the Theory of Oil Whip," by H. Poritsky, Trans. ASME, vol. 75, 1953, pp. 1153–1161.

6 "Note on Oil Whip," by Oscar Pinkus, *Journal of Applied Mechanics*, Trans. ASME, vol. 75, 1953, pp. 450–451.

7 "Analysis and Lubrication of Bearings," by M. C. Shaw and E. F. Macks, McGraw-Hill Book Company, Inc., New York, N. Y., 1949.

8 Discussion of "The Calculated Performance of Dynamically Loaded Sleeve Bearings—III," by J. T. Burwell, *Journal of Applied Mechanics*, Trans. ASME, vol. 74, 1952, pp. 239–242.

9 "The Hydrodynamic Lubrication of Cyclically Loaded Bearings," by E. M. Simons, Trans. ASME, vol. 72, 1950, pp. 805–816.

10 "Kritische Drehzahlen als Folge der Nachgiebigkeit des Schmiermittels im Lager V.D.I.," by Charles Hummel, Forschungsarbeit, 1926, heft 287.

11 "The Influence of Oil-Film Journal Bearings on the Stability of Rotating Machines," by A. C. Hagg, *Journal of Applied Mechanics*, Trans. ASME, vol. 68, 1946, pp. A-211–A-220.

12 "Some Vibration Aspects of Lubrication," by A. C. Hagg, *Lubrication Engineering*, vol. 4, August, 1948, pp. 166–169.

13 "Oil Whip of Flexible Rotors," by A. C. Hagg and P. C. Warner, Trans. ASME, vol. 75, 1953, pp. 1339–1344.

14 "Whirling Balanced Shafts," by B. L. Newkirk, Third International Congress of Applied Mechanics, Stockholm, Sweden, August, 1930.

15 "Oil-Film Whirl—A Non-Whirling Bearing," by B. L. Newkirk and L. P. Grobel, Trans. ASME, vol. 56, 1934, pp. 607–615.

16 "Instability of Oil Films and More Stable Bearings," by B. L Newkirk. General Discussion on Lubrication and Lubricants Group I. Journal and Thrust Bearings. The Institution of Mechanical Engineers, London, England, Oct. 13, 1937.

Discussion

F. T. Barwell.[5] There exists a great deal of confusion regarding oil-film whirl and this paper will be of great value in leading to an appreciation of the fact that there are at least three separate oscillatory phenomena associated with bearings under nominal constant load, namely, response to out-of-balance loading which leads to rotation of the journal center at the shaft speed, an inherent instability of lightly loaded bearings, leading to rotation of the journal at half shaft speed and, most important of all, the reaction between vibration of the rotor and the bearing which is the main subject of the present paper.

It is normally believed that this phenomenon does not occur on bearings which operate at high eccentricity and Table 1 of the paper adds support to this view. However, as the authors point out, with the heavier rotor of Table 3, instability occurs at very high eccentricities indeed. The oil-feed pressure used in the experiments was only 11 in. head but, taking bearing 38 A-C as an example, it is possible to infer from Ocvirk's work that a feed pressure of 41 psi would be required to provide the theoretical oil flow. Do the authors consider that the results of Table 3 might be affected by oil starvation and might not therefore be truly representative of practical bearings?

Oscar Pinkus.[6] The authors are to be commended for pointing out the need for an adequate and uniform set of definitions in the field of bearing instability. The identification of one form of vibration with another and the simultaneous treatment of several forms of instability caused and continues to cause wasteful argument and erroneous interpretation of the phenomena involved. There are known at present at least three forms of instability: (*a*) Unbalance which is a vibration resulting from the unbalance of the shaft which is serious only at resonant speeds. Its frequency is equal to the speed of rotation. (*b*) Oil whip or oil whirl which is a severe vibration encountered only at speeds above twice the shaft critical. It extends over a wide range of speeds and is characterized by a constant frequency, namely, that of the first critical of the shaft. (*c*) Unnamed—this third kind of instability may occur at any speed and vibrates at a frequency equal to one half of the rotating speed. The first form of vibration is known to be independent of the two others and there is little information on how much affinity there may be between the other two.

The paper points out that whip occurred also at high eccentricities, a fact corroborated by tests run by the writer. This is worth emphasizing as it has been a recent trend to assume that no oil whip occurs at high eccentricities. The pronounced and varied effect of L/D and C/D ratios on the onset of whip points out the extreme complexity of the problem which is too often overlooked in attempts to obtain a mathematical treatment of the subject.

The writer found himself at constant variance with the authors' results regarding the effect of viscosity on oil whip. It seems to the writer a reasonable thing to assume that, since damping is a stabilizing factor, a high-viscosity oil would be expected to enhance shaft stability. The difficulty in ascertaining the effect of viscosity lies partially in the fact that the eccentricity is also affected by a change of viscosity and thus the variable is not isolated. It is also possible that some dimensionless group, still to be derived, but surely containing viscosity, determines a range of stability. In that case, either an increase or decrease in viscosity may cause oil whip.

H. Poritsky.[7] The careful and painstaking work described in this paper is a valuable addition to the existing information in regard to oil whirl and rotor instability arising from it. In a general sort of way the conclusions can be summarized by the statement that oil-whirl instability sets in when ω/ω_c, the ratio of the rotational speed to the critical speed, exceeds 2; the range of stable speeds beyond $\omega/\omega_c = 2$ can be enlarged by increasing the clearance, decreasing the ratio of length of the bearing to its diameter, increasing the projected unit area load on the bearing, and decreasing the oil viscosity. All of these changes, in the directions indicated, tend to increase the eccentricity of the onset of instability. However, no clear-cut formulation in terms of the eccentricity or in terms of the other variables has been obtained by the authors. Thus, while the results of this paper are of great practical aid, the basic causes of oil whirl and the exact conditions under which it will or will not occur still remain to be formulated.

In the writer's treatment of oil whirl (reference 5 of the authors' paper), he tried to present a simple theory in terms of an elastic radial spring constant of the oil film, utilizing for the other proper-

[5] Lubrication and Wear Division, Mechanical Engineering Research Laboratory, Thorntonhall, Glasgow, Scotland.

[6] Thomson Laboratory, General Electric Company, W. Lynn, Mass. Assoc. Mem. ASME.

[7] Consulting Engineer, General Engineering Laboratory, General Electric Company, Schenectady, N. Y. Mem. ASME.

ties of the film expressions derived from the Harrison equations for long bearings. It would be of interest to find out whether by varying the radial spring constant of the film as a function of frequency and eccentricity, possibly also by modifying the Harrison terms properly, one would arrive at some consistent set of criteria which would actually explain the results obtained.

It is the writer's opinion that the basic problem ahead in the study of oil whirl is determination of the forces on the journal exerted by the oil film, and the way these forces depend upon the journal dimensions, the clearance, the eccentricity of the journal, the speed of rotation, the instantaneous velocity components of the journal center; a further variable may be the angle between the oil-inlet location and the plane of displacement of the center. Once the dependence of the forces on these quantities is obtained, either from theory or by measurements, the prediction of regions of oil whirl should be possible from the dynamical equations for the rotor motion. Until an understanding of these forces is obtained, a great deal of the work on oil whirl will cover by necessity only particular ranges of the various variables involved without yielding an over-all criterion which can be used effectively in predicting where rotor instability due to oil whirl is likely to occur, and how it can best be avoided.

Authors' Closure

Dr. Barwell has raised an important point. The flow through the bearing was measured in many of the runs. For example, in run 43 G-I the flow was 1900 cc per min. A recent measure of the flow through the orifice with oil of the same viscosity and with no journal in the bearing gave a flow of 5.7 liters per min. This was the maximum output of the pump, and it did not maintain the 11-in. head in the standpipe. It would seem that the head and the size of the tube were ample to supply all of the oil that the bearings would take at the top point, without pressure.

There were undoubtedly regions, in the bearings, of subatmospheric pressure or where the oil contained bubbles; however, it is believed that the same is true of bearings in service.

Some tests were made with a starved bearing and it was found that under the conditions of 46 D-F but with oil supply from 60 drops per min to 166 cc per min the resonant oil-film whirl did not develop at speeds up to 7500 rpm. The oil flow measured in the 46 D-F runs was 928 cc per min.

An oil pressure of 41 psi, applied to any point in the unloaded region of the bearing, would undoubtedly have had an important effect on the whirl-impending speeds.

Mr. Pinkus has mentioned some of the divergencies between his results and those reported in the foregoing. An essential difference lies in the higher critical speeds and lighter unit-bearing loads of his apparatus. The elasticity of the oil film may play a major role in his field. In this investigation the shaft is the major elastic factor.

It appears now that the indications of journal instability found by Harrison, Robertson, and others, and the two whirl phenomena referred to herein as "substantially different" may all be reconciled if due consideration is given to the elasticity of the rotor and the oil film. In the half-running-speed whirl the oil film is dominant and greater viscosity of the oil film may have a damping effect. On the other hand, with the resonant whirl the shaft elasticity determines the frequency and the film furnishes a stimulus which may be greater with oil of greater viscosity.

This view is developed in a note to be presented at the next ASME-ASLE Conference on Lubrication.

Dr. Poritsky's point is well taken in that the experimental approach has its very definite limitations. The results reported here and elsewhere should furnish check points for any analysis that may lead to a criterion for stability. It would seem that the elastic properties of the rotor must be included in any analysis that will explain the observations reported here.

The indications for design are not so simple as one would wish them to be. Examination of the tables and figures shows that the ratio $\omega/\omega_c = 2$ is not exceeded very much without running on very thin oil films, except in the case represented by the highest curve of Fig. 3. This is the case of rotor number one with a bearing having an L/D ratio of $^1/_2$ a unit load of 48 psi, and a clearance ratio of four parts per thousand. The designer must understand that there are optimum values of these variables, especially the unit loading, and that the ranges for satisfactory design are narrow.

Establishing the facts that with this apparatus the resonant whirl starts and continues at the natural frequency of the rotor for transverse vibration, that this varies slightly with running conditions, and that the half-running-frequency whirl occurs sometimes simultaneously with the resonant whirl, must contribute toward clarification of rotor disturbances due to action of the fluid films in the bearings.

While it is not exactly germane to this investigation, in view of statements appearing in the discussion, it should be pointed out that there are various causes of disturbances of rotating shafts other than unbalance and fluid-film effects. Among these are: Internal friction of the rotor, out-of-round of a journal, shaft-packing leakage, and shaft rubbing.

Load Capacity and Time Relations for Squeeze Films

By F. R. ARCHIBALD,[1] CAMBRIDGE, MASS.

The transient load capacity of a fluid film between two surfaces having a relative normal velocity is often of importance and the term squeeze film applied to such a situation is an accurate and descriptive one. Although the subject is far from new in the literature of lubrication it does not seem to have been treated very fully by any one author. It is the purpose here to show how the transient load capacity and the time to change thickness of various film configurations can be calculated and to catalog a number of results of practical interest some of which do not seem to have been given previously. There is also an interesting analogy to the torsion problem and to the problem of the flow of viscous fluids in prismatical tubes which is described.

Introduction

THE differential equation for the pressure generated in hydrodynamic lubrication, as developed by Osborne Reynolds (1),[2] includes a term for the relative velocity of approach of the lubricated surfaces in addition to the term for the relative sliding velocity. Reynolds actually considered the case of an elliptical flat plate separated by a fluid film from another flat plate of indefinite size and having a relative velocity of approach but with no relative sliding velocity. He solves this problem completely for the condition of parallel approach.

The term "squeeze film," applied to this situation of two lubricated surfaces having a relative velocity of approach, is a modern one and is due to A. F. Underwood (2). The first work on the problem seems to have been done by the physicist, Josef Stefan, in 1874 (3). Stefan solved the problem for a circular flat plate.

The squeeze-film load capacity is often of importance but it does not seem to have been treated very fully by any one author. It is the purpose here to show how this load capacity can be calculated and to catalog a number of results of practical interest, some of which do not seem to have been given previously. There is also an interesting analogy to certain other important physical problems which will be mentioned.

Squeeze-Film Problem

For problems where there is symmetry, and in which the pressure in the film is a function of only one variable, the elementary formula for viscous flow through a wide slot can be used to advantage in the squeeze-film problem. This method of solution has been employed by Fuller (4) and Purday (5).

Referring to Fig. 1, it is easily shown that the quantity of fluid flowing through the slot of length l, width b, and thickness h is given by

$$q = \frac{\Delta p b h^3}{12 \mu l} \quad \ldots\ldots [1]$$

Δp is the pressure drop in length l; hence $\Delta p/l$ is the pressure gradient; μ is the fluid viscosity.

Consider first the problem of the load capacity of the film in a journal bearing. Fig. 2 shows a journal bearing displaced from

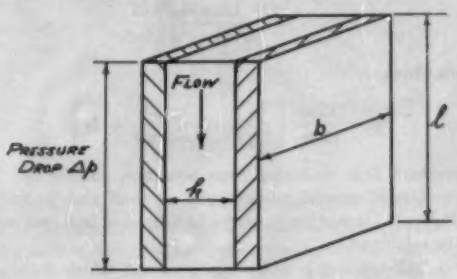


Fig. 1

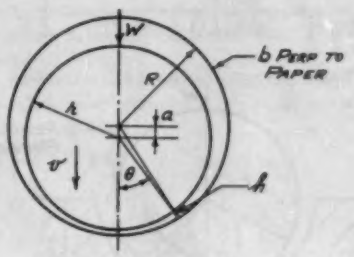


Fig. 2

its central position by a distance a. From Fig. 2 the following approximate relations can be written

$$h + r + a \cos \theta = R$$

Let

$$R - r = mr$$

Then

$$h + a \cos \theta = mr$$

Define eccentricity as

$$c = \frac{a}{R - r} = \frac{a}{mr}$$

Then

$$h = mr(1 - c \cos \theta) \quad \ldots\ldots [2]$$

---

[1] Analyst, Mechanical Division, A. D. Little, Inc. Mem. ASME.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Research Committee on Lubrication and presented at a joint session of the Lubrication Activity and Railroad Division at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 4, 1954. Paper No. 54—A-50.

The journal of Fig. 2 is being forced downward at some speed v by the load W. If inertial effects and end leakage are neglected, relations for pressure and load can be obtained as follows: The quantity of fluid passing by a radial section at θ can be written down from the equation for flow through a slot as follows

$$q = -\frac{bh^3}{12\mu}\frac{dp}{rd\theta}$$

The negative sign arises because the pressure is decreasing as θ is increasing. If h from Equation [2] is substituted

$$q = -\frac{bm^3r^3(1 - c\cos\theta)^3 dp}{12\,\mu r d\theta} \quad \ldots\ldots [3]$$

Another relation for q can be written on the basis of displacement, viz.

$$q = vbr\sin\theta \quad \ldots\ldots [4]$$

Combining Equations [3] and [4] leads to

$$dp = -\frac{12\mu v\sin\theta\, d\theta}{m^3r(1 - c\cos\theta)^3} \quad \ldots\ldots [5]$$

From which

$$p = \frac{12\mu v}{m^3 r}\left[\frac{1}{2c(1 - c\cos\theta)^2} + D\right] \quad \ldots\ldots [6]$$

The constant D is evaluated from boundary conditions. There are two cases of especial interest, viz., the half and the full bearing. Fuller (4) treats the case of a half-bearing but does not give the time relationship.

For a half-bearing $p = 0$ when $\theta = \pi/2$ which makes $D = -1/2c$ and

$$p = \frac{6\mu v}{m^3 rc}\left[\frac{1}{(1 - c\cos\theta)^2} - 1\right] \quad \ldots\ldots [7]$$

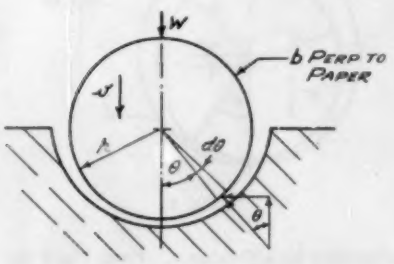


FIG. 3

Referring to Fig. 3, the load W is obtained as follows

$$W = 2br\int_0^{\pi/2} p\cos\theta d\theta$$

$$\therefore W = 2br\frac{6\mu v}{m^3 rc}\int_0^{\pi/2}\left[\frac{1}{(1 - c\cos\theta)^2} - 1\right]\cos\theta\, d\theta$$

Upon integrating and placing in the limits

$$W = \frac{12\mu vb}{m^3}\left[\frac{c}{1 - c^2} + \frac{2}{(1 - c^2)^{3/2}}\tan^{-1}\sqrt{\frac{1 + c}{1 - c}}\right] \quad \ldots [8]$$

It is clear from Fig. 2 that $v = da/dt$ and previously c has been defined as $c = a/mr$, hence

$$v = mr\frac{dc}{dt}$$

$$\therefore W = \frac{12\mu bmr}{m^3}\frac{dc}{dt}\left[\frac{c}{1 - c^2} + \frac{2}{(1 - c^2)^{3/2}}\tan^{-1}\sqrt{\frac{1 + c}{1 - c}}\right]$$

$$\therefore \int_{t_1}^{t_2} dt = \frac{12\mu br}{m^3 W}\int_{c_1}^{c_2}\left[\frac{c}{1 - c^2} + \frac{2}{(1 - c^2)^{3/2}}\tan^{-1}\sqrt{\frac{1 + c}{1 - c}}\right] dc$$

The first term of the integrand is readily integrated; the second term can be integrated by parts. Substitution of the limits in the integral gives

$$\Delta t = \frac{24\mu br}{m^3 W}\left[\frac{c_2}{\sqrt{1 - c_2^2}}\tan^{-1}\sqrt{\frac{1 + c_2}{1 - c_2}} - \frac{c_1}{\sqrt{1 - c_1^2}}\tan^{-1}\sqrt{\frac{1 + c_1}{1 - c_1}}\right] \quad \ldots\ldots [9]$$

where Δt = the time interval $(t_2 - t_1)$.

For a full bearing it is assumed that the pressure is zero on the load side of the journal, i.e., at $\theta = \pi$, and this makes $D = -1/2c(1 + c)^2$ in Equation [6]. The pressure at any point in a full bearing is then given by

$$p = \frac{6\mu v}{m^3 rc}\left[\frac{1}{(1 - c\cos\theta)^3} - \frac{1}{(1 + c)^2}\right] \quad \ldots\ldots [10]$$

The load W is found in the same manner as for the half-bearing and

$$W = 2br\frac{6\mu v}{m^3 rc}\int_0^{\pi/2}\left[\frac{1}{(1 - c\cos\theta)^3} - \frac{1}{(1 + c)^3}\right]\cos\theta d\theta$$

which upon integration and substitution of limits gives

$$W = \frac{12\pi\mu vb}{m^3(1 - c^2)^{3/2}} \quad \ldots\ldots [11]$$

Again making the substitution

$$v = mr\frac{dc}{dt}$$

$$W = \frac{12\pi\mu b}{m^3(1 - c^2)^{3/2}}mr\frac{dc}{dt}$$

$$\int_{t_1}^{t_2} dt = \frac{12\pi\mu br}{m^3 W}\int_{c_1}^{c_2}\frac{dc}{(1 - c^2)^{3/2}}$$

$$\therefore \Delta t = \frac{12\pi\mu br}{m^3 W}\left[\frac{c_2}{\sqrt{1 - c_2^2}} - \frac{c_1}{\sqrt{1 - c_1^2}}\right] \quad \ldots\ldots [12]$$

It should be borne in mind that the foregoing relations for a journal bearing have been derived on the assumption of no end leakage, and they must, therefore, be used with considerable reserve. Because of the end leakage, there is a good possibility of film failure on the top side of the full bearing. Such film failure is not as likely in a half-bearing and so the results given for this case should be more reliable. If significant film failure should occur in the full bearing, it would tend toward the condition of the half-bearing and it might very well be better to use the half-bearing formulas even for a full bearing.

Case of a Spherical Bearing

The case of a spherical bearing is easily solved in the same way as for the journal bearing, and since there is no side leakage the result should be in good agreement with experiment. The film geometry in diametral planes is the same as for the journal bearing and, therefore, the film thickness h is again given by Equation [2]. The case of a hemispherical seat will be considered.

By using Equation [1], the quantity of fluid, passing a conical

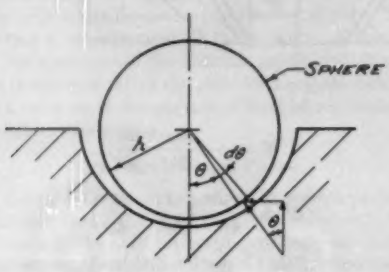


Fig. 4

surface defined by the angle θ in Fig. 4, is given by

$$q = -\frac{2\pi r \sin\theta m^3 r^3 (1 - c\cos\theta)^3 dp}{12\mu r d\theta} \quad \ldots\ldots [13]$$

Another relation for q is obtained from displacement

$$q = \pi r^2 v \sin^2\theta \quad \ldots\ldots [14]$$

Combining Equations [13] and [14]

$$dp = -\frac{6\mu v \sin\theta d\theta}{m^3 r (1 - c\cos\theta)^3} \quad \ldots\ldots [15]$$

From which

$$p = -\frac{6\mu v}{m^3 r}\left[-\frac{1}{2c(1 - c\cos\theta)^2} + A\right] \quad \ldots\ldots [16]$$

When $\theta = \pi/2$, $p = 0$ which makes $A = 1/2c$

$$\therefore p = \frac{3\mu v}{m^3 r c}\left[\frac{1}{(1 - c\cos\theta)^2} - 1\right] \quad \ldots\ldots [17]$$

From the figure the load W is given by

$$W = 2\pi r^2 \int_0^{\pi/2} p \sin\theta\cos\theta\, d\theta$$

$$= 2\pi r^2 \frac{3\mu v}{m^3 r c}\int_0^{\pi/2}\left[\frac{1}{(1 - c\cos\theta)^2} - 1\right]\sin\theta\cos\theta d\theta$$

From which

$$W = \frac{6\pi\mu v r}{m^3}\left[\frac{1}{c^3}\log(1 - c) + \frac{1}{c^2(1 - c)} - \frac{1}{2c}\right] \quad \ldots [18]$$

(As $c \to 0$ the bracketed quantity has a limit of 2/3.)

As in the case of the journal bearing $v = mrdc/dt$ and hence

$$\int_{t_1}^{t_2} dt = \frac{6\pi\mu r^2}{m^2 W}\int_{c_1}^{c_2}\left[\frac{1}{c^3}\log(1 - c) + \frac{1}{c^2(1 - c)} - \frac{1}{2c}\right]dc$$

$$\therefore \Delta t = \frac{3\pi\mu r^2}{m^2 W}\left[-\left(\frac{1 - c}{c}\right) - \left(\frac{1 + c^2}{c^2}\right)\log(1 - c)\right]_{c_1}^{c_2} \quad \ldots [19]$$

(As $c \to 0$ the bracketed quantity has a limit of 3/2.)

$$\Delta t = \frac{3\pi\mu r^2}{m^2 W}\left[\frac{c_2 - c_1}{c_1 c_2} + \left(\frac{1 + c_1^2}{c_1^2}\right)\log(1 - c_1) - \left(\frac{1 + c_2^2}{c_2^2}\right)\log(1 - c_2)\right] \quad \ldots [20]$$

The "Workshop Viscometer"

The motion of a ball in a spherical seat has been applied by Michell (6) to the determination of viscosity in his workshop viscometer. This viscometer is shown in cross section in Fig. 5.

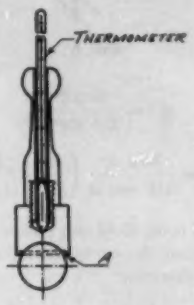


Fig. 5

The ball is 1 in. in diam of stainless steel (approximately 0.148 lb). A minimum clearance between ball and cup of the order of 0.0004 in. is maintained by three very small projections spaced 120 deg apart on the surface of the cup. A few drops of oil to be tested are placed in the cup. The ball is then dropped into place and most of the oil is displaced to the recess at A. The instrument is inverted and simultaneously a time reading is taken. The time required for the ball to become free is observed. This is a measure of the viscosity.

By similar methods it can be shown that the time for a ball of weight W to fall from an eccentricity c_1 to an eccentricity c_2 is given by

$$\Delta t = \frac{3\pi\mu r^2}{m^2 W}\left[\left(1 + \frac{1}{c_2^2}\right)\log\left(\frac{1 - c_2}{1 - c_2\cos\alpha}\right) - \left(1 + \frac{1}{c_1^2}\right)\log\left(\frac{1 - c_1}{1 - c_1\cos\alpha}\right) + \frac{\sin^2\alpha}{1 - c_2\cos\alpha} - \frac{\sin^2\alpha}{1 - c_1\cos\alpha} + \left(\frac{1}{c_2} - \frac{1}{c_1}\right)(1 - \cos\alpha)\right] \quad \ldots [21]$$

where α is half the angle subtended by the spherical seat.

It can be shown that this time interval has the following limit for the case when the ball starts from the concentric position ($c = 0$), and goes to a very large eccentricity ($c \to -\infty$)

$$\Delta t = \frac{3\pi\mu r^2}{m^2 W}\left(\log\sec\alpha - \frac{1}{2}\sin^2\alpha\right) \quad \ldots\ldots [22]$$

For Michell's instrument $r = 0.5$ in., $W = 0.148$ lb. Assuming the clearance of 0.0004 in. is a concentric clearance, then $m = 0.0008$ and taking $\alpha = 45$ deg

$$\mu = 4.16 \times 10^{-7}\Delta t \quad \text{lb sec/in.}^2 \text{ (reyns)}$$

Conical Bearing

Another problem of practical interest is the settling time of a conical bearing. Two cases arise, viz., the case of the complete cone as in Fig. 6, and the case of the truncated cone as in Fig. 7.

For the complete cone

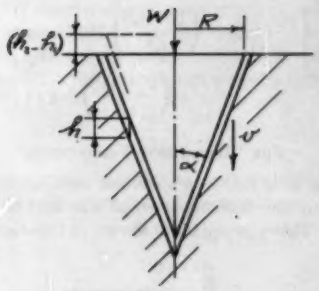


FIG. 6

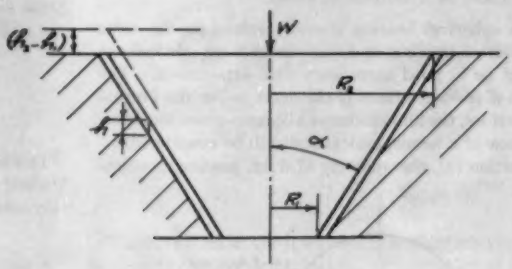


FIG. 7

$$W = \frac{3\pi\mu R^4 v}{2h^3 \sin^4 \alpha} \quad [23]$$

$$\Delta t = \frac{3\pi\mu R^4}{4W \sin^4 \alpha}\left(\frac{1}{h_2^2} - \frac{1}{h_1^2}\right) \quad [24]$$

With the truncated cone, fluid can escape from both ends and the formulas for W and Δt are quite different from Equations [23] and [24]. For this case

$$W = \frac{3\pi\mu v}{2h^3 \sin^4 \alpha}\left[R_2^4 - R_1^4 - \frac{(R_2^2 - R_1^2)^2}{\log \frac{R_2}{R_1}}\right] \quad [25]$$

$$\Delta t = \frac{3\pi\mu}{4W \sin^4 \alpha}\left[R_2^4 - R_1^4 - \frac{(R_2^2 - R_1^2)^2}{\log \frac{R_2}{R_1}}\right]\left(\frac{1}{h_2^2} - \frac{1}{h_1^2}\right) \quad [26]$$

CASE OF TWO LUBRICATED SURFACES

Problems of squeeze films, in which the film is between two parallel flat plates, can be solved for a large number of plate shapes. It has been mentioned that Reynolds basic differential equation for hydrodynamic lubrication contained a term for the velocity of approach of the two lubricated surfaces. This equation is as follows

$$\frac{\partial}{\partial x}\left(h^3 \frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial z}\left(h^3 \frac{\partial p}{\partial z}\right) = 6\mu\left(U \frac{dh}{dx} + 2v\right) \quad [27]$$

In this equation U is the relative sliding velocity of the surfaces and v is the velocity of approach. If U is zero and the plates are parallel, this equation reduces to

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} = \frac{12\mu v}{h^3} \quad [28]$$

$$v = \frac{dh}{dt} \quad [29]$$

(Equation [28] is derived from first principles in the Appendix.)

Reynolds results (1) for the case of an elliptical plate have been alluded to and are

$$W = -\frac{3\pi\mu a^3 b^3}{(a^2 + b^2)h^3}\frac{dh}{dt} \quad [30]$$

$$\Delta t = \frac{3\pi\mu a^3 b^3}{2(a^3 + b^3)W}\left(\frac{1}{h_2^2} - \frac{1}{h_1^2}\right) \quad [31]$$

The results for a circular plate are easily obtained by letting $a = b = r$ and

$$W = -\frac{3\pi\mu r^4}{2h^3}\frac{dh}{dt} \quad [32]$$

$$\Delta t = \frac{3\pi\mu r^4}{4W}\left(\frac{1}{h_2^2} - \frac{1}{h_1^2}\right) \quad [33]$$

It will be observed that Equation [28] has the same form as the well-known equation for the stress function in the torsion problem in the theory of elasticity (7)

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = -2G\theta \quad [34]$$

An equation of the same form must be satisfied by the velocity of a viscous fluid flowing in a prismatical tube (8)

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = -\frac{1}{\mu}\frac{dp}{dz} \quad [35]$$

The analogy here could be quite useful particularly in cases where the region is multiply connected. In the case of multiply connected regions in the torsion problem, the function ϕ takes on constant, but generally unequal, values on each boundary. It is usual to choose the value of ϕ on the outside boundary as zero. To produce the analogous condition in the squeeze-film problem,

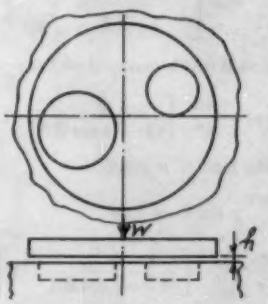


FIG. 8

the bottom plate, say, would have deep depressions in it filled with fluid as in Fig. 8.

The pressure in these depressions would be sensibly constant because of the negligible shear rates in the fluid in these regions. There is some question regarding the accuracy because of the distance required in a passage to develop Poiseuille flow. Some reassurance is obtained here from a test on a circular plate with a concentric hole. Experiment and theory were found in close agreement. If this may be taken as settling the question of error due to the establishment of the flow regime, there still remains a question on the constancy of the pressure over the depressed regions when flow passes into and out of these regions as it must in a situation like that of Fig. 8.

The corresponding problem in fluid flow would arise in a case where the flow took place through a tube in which there were prismatical rods running parallel to the tube elements. In this case the velocity w of the fluid is zero on every boundary, and if Fig. 8 is again used, the depressions would have to be kept empty by draining at the bottom. In this way the fluid flows to all boundaries where the pressure is zero and the analogy is complete.

Unfortunately, in the case of the torsion analogy, there does not seem to be a very convenient way of determining the stress (though the author has devised a procedure using dye streams), which involves a measurement of the gradient of the pressure function. But a measure of the torsional rigidity is obtained quite easily from the time of fall of the plate between two heights, and this is also a measure of the amount of fluid which would flow in a tube of similar cross section.

The Rectangular Plate

Because the problem is analogous to the torsion problem, the results available from the theory of elasticity can be used directly. Probably the most interesting of these results, from the practical standpoint, is the case of a rectangular plate for which

$$W = -\frac{192\mu a^3}{\pi^5 h^3}\frac{dh}{dt}\left[\frac{\pi^4 b}{192} - \frac{a}{\pi}\sum_{n=1,3,5\ldots}^{\infty}\frac{1}{n^5}\tanh\frac{n\pi b}{2a}\right] \quad [36]$$

$$\Delta t = \frac{96\mu a^3}{\pi^5 W}\left(\frac{1}{h_2^3} - \frac{1}{h_1^3}\right)\left[\frac{\pi^4 b}{192} - \frac{a}{\pi}\sum_{n=1,3,5\ldots}^{\infty}\frac{1}{n^5}\tanh\frac{n\pi b}{2a}\right] \quad [37]$$

where a and b are the lengths of the sides of the rectangle. If the rectangle is very long compared to its width, these formulas become approximately

$$W = -\frac{\mu a^3 b}{h^3}\frac{dh}{dt} \quad [38]$$

$$\Delta t = \frac{\mu a^3 b}{2W}\left(\frac{1}{h_2^3} - \frac{1}{h_1^3}\right) \quad [39]$$

In Equations [38] and [39], $b \gg a$. In Equations [36] and [37] it does not matter which one is taken as the larger, though for computation it is more convenient to choose b as the large dimension because of the way the hyperbolic tangent behaves for large values of its argument.

Because a thrust surface is often broken by radial grooves, to admit lubricant, the squeeze film for a region bounded by two concentric circles and two radii is of practical interest. A closely related case, that of the sector region of radius a, is also of interest. The values for W and Δt for these two regions are as follows:

For the thrust surface bounded by two concentric circles of inside radius r_1 and outside radius r_2 and subtending an angle α at the center

$$W = -\frac{96\mu\alpha}{\pi^2 h^3}\frac{dh}{dt}\sum_{n=1,3,5\ldots}^{\infty}\left[\frac{r_2^4 - r_1^4}{4n^2\left(2 - \frac{n\pi}{\alpha}\right)^2} - \frac{2\pi}{n\alpha}\frac{1}{\left(4 - \frac{n^2\pi^2}{\alpha^2}\right)^2}\frac{\left(r_2^{2+\frac{n\pi}{\alpha}} - r_1^{2+\frac{n\pi}{\alpha}}\right)^2}{r_2^{\frac{2n\pi}{\alpha}} - r_1^{\frac{2n\pi}{\alpha}}}\right] \quad [40]$$

$$\Delta t = \frac{48\mu\alpha}{\pi^2 W}\left(\frac{1}{h_2^3} - \frac{1}{h_1^3}\right)\sum_{n=1,3,5\ldots}^{\infty}\left[\frac{r_2^4 - r_1^4}{4n^2\left(2 - \frac{n\pi}{\alpha}\right)^2} - \frac{2\pi}{n\alpha}\frac{1}{\left(4 - \frac{n^2\pi^2}{\alpha^2}\right)^2}\frac{\left(r_2^{2+\frac{n\pi}{\alpha}} - r_1^{2+\frac{n\pi}{\alpha}}\right)^2}{r_2^{\frac{2n\pi}{\alpha}} - r_1^{\frac{2n\pi}{\alpha}}}\right] \quad [41]$$

For the sector region of radius a and included angle α

$$W = -\frac{3\mu a^4}{4h^3}\frac{dh}{dt}\left[\alpha - \tan\alpha + \frac{128\alpha^4}{\pi^6}\sum_{n=1,3,5\ldots}^{\infty}\frac{1}{n^2\left(n + \frac{2\alpha}{\pi}\right)^2\left(n - \frac{2\alpha}{\pi}\right)}\right] \quad [42]$$

$$\Delta t = \frac{3\mu a^4}{8W}\left(\frac{1}{h_2^3} - \frac{1}{h_1^3}\right)\left[\alpha - \tan\alpha + \frac{128\alpha^4}{\pi^6}\sum_{n=1,3,5\ldots}^{\infty}\frac{1}{n^2\left(n + \frac{2\alpha}{\pi}\right)^2\left(n - \frac{2\alpha}{\pi}\right)}\right] \quad [43]$$

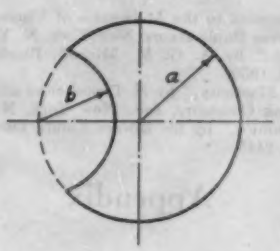


Fig. 9

For the shape shown in Fig. 9, the formulas for W and Δt are

$$W = -\frac{3\mu}{2h^3}\frac{dh}{dt}\left[(2a^4 - 4a^2b^2 - b^4)\cos^{-1}\frac{b}{2a} + \frac{b}{2}\left(a^2 + \frac{7}{2}b^2\right)\sqrt{4a^2 - b^2}\right] \quad [44]$$

$$\Delta t = \frac{3\mu}{4W}\left(\frac{1}{h_2^3} - \frac{1}{h_1^3}\right)\left[(2a^4 - 4a^2b^2 - b^4)\cos^{-1}\frac{b}{2a} + \frac{b}{2}\left(a^2 + \frac{7}{2}b^2\right)\sqrt{4a^2 - b^2}\right] \quad [45]$$

If $b \to 0$, these formulas degenerate to the case of a circular plate.

Circular Plate With Concentric Hole

As a concluding example, the case of a circular plate having a concentric hole will be considered. Two cases of practical interest can arise. Both can be solved by the elementary method using Equation [1]. If, as in Fig. 8, the hole does not allow any escape of fluid and the entire flow takes place over the outside boundary

$$W = -\frac{3\pi\mu}{2h^3}\frac{dh}{dt}(r_2^2 - r_1^2) \quad [46]$$

$$\Delta t = \frac{3\pi\mu}{4W}(r_2^2 - r_1^2)\left(\frac{1}{h_2^2} - \frac{1}{h_1^2}\right) \quad [47]$$

If the flow can take place over the inside boundary as well as the outside boundary

$$W = -\frac{3\pi\mu}{2h^3}\frac{dh}{dt}\left[r_2^4 - r_1^4 - \frac{(r_2^2 - r_1^2)^2}{\log\frac{r_2}{r_1}}\right] \quad [48]$$

$$\Delta t = \frac{3\pi\mu}{4W}\left(\frac{1}{h_2^2} - \frac{1}{h_1^2}\right)\left[r_2^4 - r_1^4 - \frac{(r_2^2 - r_1^2)^2}{\log\frac{r_2}{r_1}}\right] \quad [49]$$

In a wet-disk clutch, the time to seat can be estimated by using Equation [49].

BIBLIOGRAPHY

1 "On the Theory of Lubrication," by O. Reynolds, Philosophical Transactions of the Royal Society of London, England, part 1, vol. 177, 1886, pp. 157–234.

2 "Rotating Load Bearings—A New Concept of Operation and a Frictionless Support," by A. F. Underwood, ASME Paper No. 44—A-29; *Automotive Industries*, vol. 92, no. 11, 1945, p. 26.

3 "Versuche über scheinbare Adhesion," by Josef Stefan, *Berichte Kaiserlich Akademie d. Wissenschaften*, vol. 69, 1874, pp. 713–734.

4 "Hydrostatic Lubrication—Part IV, Oil Cushions," by D. D. Fuller, *Machine Design*, vol. 19, September, 1947, pp. 127–131, 188–189.

5 "An Introduction to the Mechanics of Viscous Flow," by H. F. P. Purday, Dover Publications, New York, N. Y., 1949.

6 "Lubrication," by A. G. M. Michell, Blackie & Son, Ltd., London, England, 1950.

7 "Theory of Elasticity," by S. Timoshenko and J. N. Goodier, McGraw-Hill Book Company, Inc., New York, N. Y., 1951.

8 "Hydrodynamics," by Sir Horace Lamb, Dover Publications, New York, N. Y., 1945.

Appendix

The equation

$$\frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} = \frac{12\mu v}{h^3}$$

will be derived from first principles.

Consider an element in a fluid between a pair of flat parallel plates being pulled apart with a velocity v. The pressures and shear stress on the faces of such an element are shown in Fig 10. It is assumed that the pressure is constant through the thickness of the film.

For the condition of statical equilibrium in the x-direction, the following equation is written

$$p\,dy\,dz - \left(p + \frac{\partial p}{\partial x}\,dx\right)dy\,dz + \left(\tau_{xy} + \frac{\partial \tau_{xy}}{\partial y}\,dy\right)dx\,dz - \tau_{xy}\,dx\,dz = 0$$

from which

$$\frac{\partial p}{\partial x} = \frac{\partial \tau_{xy}}{\partial y} \quad [50]$$

Similarly, from equilibrium in the z-direction

$$\frac{\partial p}{\partial z} = \frac{\partial \tau_{yz}}{\partial y} \quad [51]$$

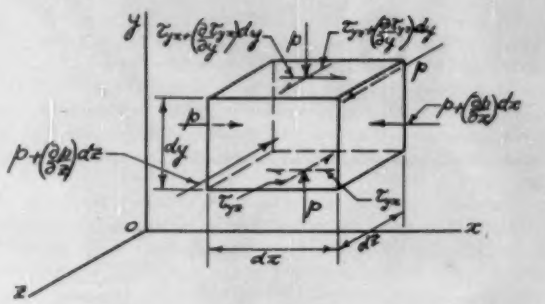


FIG. 10

Now from the relation between shear stress and velocity gradient in a viscous fluid (Newton's law)

$$\tau_{xy} = \mu\frac{\partial u}{\partial y} \quad [52]$$

Hence

$$\frac{\partial^2 u}{\partial y^2} = \frac{1}{\mu}\frac{\partial p}{\partial x} \quad [53]$$

Since p is not a function of y this equation is integrated as follows

$$\frac{\partial u}{\partial y} = \frac{1}{\mu}\frac{\partial p}{\partial x}\,y + f(x, z)$$

$$u = \frac{1}{2\mu}\frac{\partial p}{\partial x}\,y^2 + yf(x, z) + g(x, z) \quad [54]$$

The arbitrary functions $f(x, z)$ and $g(x, z)$ are evaluated from the boundary conditions:

when $y = 0$, $u = 0$

when $y = h$, $u = 0$

which give

$$f(x, z) = -\frac{h}{2\mu}\frac{\partial p}{\partial x} \text{ and } g(x, z) = 0$$

$$\therefore u = \frac{1}{2\mu}\frac{\partial p}{\partial x}\,y^2 - \frac{1}{2\mu}\frac{\partial p}{\partial x}\,hy \quad [55]$$

The rate of flow in the x-direction at any section of the film is given by

$$q_x = \int_0^h u\,dy = \int_0^h \left[\frac{1}{2\mu}\frac{\partial p}{\partial x}\,y^2 - \frac{1}{2\mu}\frac{\partial p}{\partial x}\,hy\right]dy$$

$$\therefore q_x = -\frac{h^3}{12\mu}\frac{\partial p}{\partial x} \quad [56]$$

In the same way, the rate of flow in the z-direction is found as

$$q_z = -\frac{h^3}{12\mu}\frac{\partial p}{\partial z} \quad [57]$$

Consider a small prism of fluid in the film of width dx, length

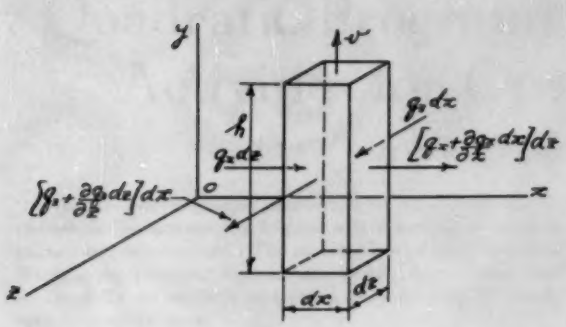


FIG. 11

dz, and height h (see Fig. 11). For an incompressible fluid, the following continuity relation can be made

$$-\left(q_x + \frac{\partial q_x}{\partial x} dx\right) dz + q_x dz - \left(q_z + \frac{\partial q_z}{\partial z} dz\right) dx + q_z dx - v\, dx\, dz = 0$$

Hence

$$\frac{\partial q_x}{\partial x} + \frac{\partial q_z}{\partial z} = -v \quad \ldots\ldots\ldots\ldots\ldots\ldots [58]$$

Substituting from Equations [56] and [57]

$$-\frac{h^3}{12\mu}\frac{\partial^2 p}{\partial x^2} - \frac{h^3}{12\mu}\frac{\partial^2 p}{\partial z^2} = -v$$

$$\therefore \frac{\partial^2 p}{\partial x^2} + \frac{\partial^2 p}{\partial z^2} = \frac{12\mu}{h^3} v \quad \ldots\ldots\ldots\ldots\ldots [59]$$

Quadratic Programming of Interdependent Activities for Optimum Performance

By LINDON E. SALINE,[1] SCHENECTADY, N. Y.

This paper develops and demonstrates a generalized procedure for programming a quadratic function to achieve optimum performance. The method is a generalization of linear programming techniques which have been used successfully to achieve optimum performance in many economic situations.

Introduction

STUDIES relating to the performance of economic systems frequently have associated with them problems of programming interdependent activities. Programming interdependent activities is scheduling the levels of operation of the activities to achieve a desired goal and to meet specified requirements. These economic systems may involve the economy of the entire nation or a small part of one business. Commonly it is desired to determine the balance between interdependent activities that gives maximum profit or minimum cost although other measures of effectiveness could be employed.

One method of arriving at a maximum profit or minimum cost (optimum performance) is to cut and try until system performance cannot be improved further. A newer method is to use the technique of "programming." Linear programming has received considerable attention in the literature[2] and has been used successfully to achieve optimum performance in many economic situations. Linear programming—as the name implies—is restrictive in its application to the optimization of linear functions. The purpose of this paper is to develop and demonstrate a method of programming a generalized quadratic function; linear programming is contained as a special case of the generalized quadratic function.

The Programming Approach

The programming approach to optimization problems is simply an iterative procedure. A solution to a set of linear constraining equations is continually improved until a certain nonlinear expression (the functional) involving the variables found in the solution to the constraining equations reaches a maximum or minimum. The nonlinear expression to be maximized or minimized is the functional and measures the effectiveness of performance of the system.

For example, in the classical transportation problem the constraining equations describe the goods available at all sending points and the goods required at all receiving points. The functional describes the cost of shipping the goods between all possible sending and receiving points. The programming problem is to schedule the flow of goods which minimizes the total cost of shipping and satisfies the requirements at the sending and receiving points.

Generally there are fewer constraining equations than unknowns and hence the constraining equations have an infinite number of possible solutions. Any solution to the constraining equations (not necessarily the solution which gives the optimum value to the functional) is called a "basis." Programming is essentially an iterative procedure which starts with any basis and proceeds to improve the basis in successive steps until a basis which optimizes the functional is found.

The method of improving the basis is to calculate the incremental change in the functional due to the addition of an incremental amount of one of the unknowns to the basis.

A positive increment[3] in the functional indicates that the functional is improved by adding the unknown into the basis. The basis, of course, must be adjusted so it satisfies the constraining equations. An optimum basis has been obtained when the increment for the functional is either zero or negative for all possible incremental additions to the basis.

Problem Statement

The generalized quadratic programming problem is to maximize (or minimize) the functional

$$C = \sum_{i=1}^{r} a_i\lambda_i + \sum_{i=1}^{r} b_i\lambda^2_i + \sum_{\substack{i=1 \\ j=i+1}}^{\substack{j=r \\ i=r-1}} c_{ij}\lambda_i\lambda_j + \sum_{i=1}^{r} \frac{d_i\lambda_i}{\lambda_i + \delta} \quad . . [1]$$

where a_i, b_i, c_{ij}, and d_i are positive or negative constants, and δ is an arbitrarily small number. The last term is essentially the constant d_i which represents a "fixed" value if $\lambda_i > 0$; differentiation and subsequent integration of the last term as written in the equation does not destroy d_i for future use. λ_i and λ_j are positive variables restricted according to linear Equations [2]

x_{i0}	λ_1	$\lambda_2 \ldots \lambda_k \ldots \lambda_m$	$\lambda_n \ldots \ldots \lambda_r$	
x_{10}	1		x_{1n}	x_{1r}
x_{20}		1	x_{2n}	x_{2r}
—				
—				
—				
—				
—				
—				
x_{k0}		1	x_{kn}	x_{kr}
—				
—				
—				
x_{m0}		1	x_{mn}	x_{mr}

. . [2]

[1] Currently in Project Analysis Section, Research Laboratory, General Electric Company. This work was done while author was in Analytical Engineering Section of Apparatus Sales Division. Assoc. Mem. ASME.

[2] The book "An Introduction to Linear Programming," by A. Charnes, W. W. Cooper, and A. Henderson (John Wiley & Sons, Inc., New York, N. Y., 1953), is an excellent reference to the subject of linear programming and contains an excellent bibliography.

Contributed by the Management Division and presented at a joint meeting with the Production Engineering Division at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, August 6, 1954. Paper No. 54—A-58.

[3] The statements which follow assume that the functional is to be maximized. If the functional is to be minimized, recall that min $f(x) = \max[-f(x)]$. Hence the rules which are developed here apply to either maximization or minimization problems.

It should be noted that if $\lambda_a = \lambda_{a+1} = \ldots = \lambda_r = 0$, then $\lambda_1 = x_{10}, \lambda_2 = x_{20}, \ldots, \lambda_m = x_{m0}$ is a solution (not necessarily the optimum) to the constraining equations. The vectors $\lambda_1 \ldots \lambda_m$ are said to form the basis. The method of arranging the equations in this fashion will be discussed subsequently. The value of the functional for the foregoing basis is simply C, Equation [1], with $r = m$.

At first glance the programming problem may appear as a problem amenable to the techniques of the calculus of variations. However, the distinction is that the unknown variables must be greater than, or equal to, zero in the programming problem, whereas the variables are not restricted in this manner in the calculus of variations problem. This restriction is one that is encountered in many business problems and one that cannot be handled by the calculus of variations except in certain special problems in which all of the variables happen to be zero or positive in the closed solution. As a consequence of the restriction that all the variables must be zero or positive, a further limitation on the calculus of variations is that it cannot, in general, handle constraining relations which are expressed as equalities.

Calculating Routine

The determination of the optimum solution as developed in Appendix 2 can be reduced to a routine schedule for either hand or machine calculations. The use of tableaux in a manner similar (but not exactly the same) to that suggested by Charnes, et al., is illustrated for a problem relating to the allocation of manufacturing facilities in Appendix 1. Computing machines can be programmed to do the calculations, but efficient calculation requires a large memory unit.

The following steps should guide the user through the sample calculating schedule presented:

1 Fill in constants associated with the functional and the constraining equations.

2 Fill in the lower part of the table by following steps 1 to 7 as shown on the schedule.

3 Determine the variable associated with the largest ΔC as calculated in Step 7.

4 Fill in the constants for the constraining equations on the second tableau by using the elimination method as indicated in Equations [8a] and [8c] of Appendix 2. The constraining equation containing the new variable introduced into the basis should occupy the first row in the tableau.

5 Fill in the constants associated with the functional. Calculate Step 1 on the tableau. The value obtained for C_{Total} in Tableau 2 should equal the C_{Total} in Tableau 1 plus the ΔC (Step 7) calculated in Tableau 1.

6 Continue these calculations until all of the ΔC are zero or negative. The values appearing in the x_{i0} column are then the optimum values which are associated with the appropriate variables in the constraining equations.

Appendix 1

Example Relating to Allocation of Manufacturing Facilities

Elements of Cost. The principal factors of cost relating to the economic allocation of manufacturing facilities are manufacturing costs, transportation costs, and inventory costs of the finished goods at the consuming point.

Manufacturing costs include material (purchase and inventory), labor, and all indirect manufacturing expenses (space rental, electric power, depreciation, and so on). Transportation costs include all charges incurred in shipping goods between the manufacturing point and the consuming point. Inventory costs for the finished goods at the consuming point include charges on investment, handling, warehousing, and obsolescence.

Constraints of Manufacturing Facilities and Goods Required

Goods are supplied by m manufacturing plants to n consuming points. The m manufacturing plants have a specified maximum capacity P, and the n consuming points require a specified amount of goods C. These constraining conditions can be given in tabular form as shown herewith.

Consuming points \ Manufacturing plants	1	2 . . . . . m	Required amount specified at each consuming point
1	λ_{11}	$\lambda_{12} \ldots \ldots \lambda_{1m}$	$C_1 = \sum_{k=1}^{m} \lambda_{1k}$
2	λ_{21}	$\lambda_{22} \ldots \ldots \lambda_{2m}$	$C_2 = \sum_{k=1}^{m} \lambda_{2k}$
,	,		,
,	,		,
,	,		,
,	,		,
,	,		,
,	,		,
n	λ_{n1}	$\lambda_{n2} \ldots \ldots \lambda_{nm}$	$C_n = \sum_{k=1}^{m} \lambda_{nk}$
Unknown total supplied by each manufacturing plant	$\sum_{i=1}^{n} \lambda_{i1}$	$\sum_{i=1}^{n} \lambda_{i2} \ldots \sum_{i=1}^{n} \lambda_{im}$	Grand total = $\sum_{i=1}^{i=n} \sum_{k=1}^{k=m} \lambda_{ik}$
Specified maximum capacity for each manufacturing plant	P_1	$P_2 \ldots \ldots P_m$	

The λ_{ik} are the unknown quantities of goods supplied by manufacturing plant k to consuming point i. In some cases, manufacturing plant q would not be allowed to supply consuming point p; then $\lambda_{pq} \equiv 0$. From the preceding table there are $(m + n)$ constraining equations and inequations and (mn—number of impossible λ_{pq}) variables.

Allocation as a Programming Problem

Programming interdependent activities is scheduling the levels of operation of the activities to achieve a desired goal and to meet specified requirements. The desired goal in an allocation problem is to achieve minimum cost or maximum profit; the specified requirements are the capabilities of the manufacturing plants and the amounts of goods required at the consuming points.

The problem is to minimize over-all cost

$$C = M(\lambda_{ik}) + T(\lambda_{ik}) + I(\lambda_{ik})$$

where M, T, and I represent the manufacturing, transportation, and inventory costs which are functions of the λ_{ik}, while satisfying constraining equations of the form

$$P_1 = \lambda_1 + \lambda_{11} + \lambda_{21} + \ldots + \lambda_{n1}$$

(the sum of the goods from manufacturing plant 1 must be less than, or equal to, P_1)

$$C_1 = \lambda_2 + \lambda_{11} + \lambda_{12} + \ldots + \lambda_{1m}$$

(the sum of the goods into consuming point 1 must be equal to C_1), where λ_1 is a slack vector which has zero unit cost and λ_2 is an artificial vector which has unit cost Z (a large unbounded number).

Numerical Example

A company having two manufacturing plants and three distributors wants to determine the optimum annual manufacturing rate to supply known rates of requirements. The annual cost of manufacturing is shown in Fig. 1 for plants 1 and 2 as a function of the total output for each plant. Transportation costs are also shown in Fig. 1 for three zones of transportation. The transportation rates applicable to each variable are indicated in a subsequent table. The inventory rates correspond to a per cent of annual goods for each variable and are determined by the reliability of the manufacturing plant, the transportation system, and the requirements forecast. The tabular description of the system is shown herewith.

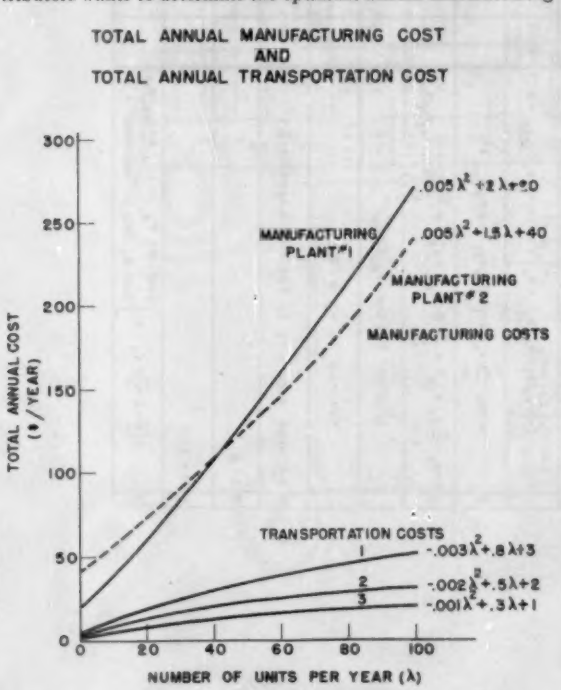


Fig. 1 Total Annual Manufacturing Cost and Total Annual Transportation Cost

Consuming plants \ Manufacturing plants	1	2	Required amount specified at each consuming point
1	λ_{11}	λ_{12}	50
2	λ_{21}	λ_{22}	60
3	λ_{31}	λ_{32}	70
Specified maximum capacity for each manufacturing plant	100	100	180

Variable	Transportation costs (Curve No. in Fig. 1)	Inventory rate (per cent annual usage)
λ_{11}	3	2
λ_{12}	1	4
λ_{21}	2	3
λ_{22}	3	2
λ_{31}	3	2
λ_{32}	2	3

Constraining Equations. The constraining equations can be written as follows

$$\begin{aligned} 50 &= \lambda_1 + \lambda_{11} + \lambda_{12} \\ 60 &= \lambda_2 + \lambda_{21} + \lambda_{22} \\ 70 &= \lambda_3 + \lambda_{31} + \lambda_{32} \\ 100 &= \lambda_4 + \lambda_{11} + \lambda_{21} + \lambda_{31} \\ 100 &= \lambda_5 + \lambda_{12} + \lambda_{22} + \lambda_{32} \end{aligned}$$

where λ_1, λ_2, and λ_3 are artificial vectors and λ_4 and λ_5 are slack vectors.

Functional. The functional expression describing cost is

$$\begin{aligned} C = {} & 0.005\,(\lambda_{11} + \lambda_{21} + \lambda_{31})^2 + 2(\lambda_{11} + \lambda_{21} + \lambda_{31}) + 20 \\ & + 0.005\,(\lambda_{12} + \lambda_{22} + \lambda_{32})^2 + 1.5\,(\lambda_{12} + \lambda_{22} + \lambda_{32}) + 40 \\ & - 0.001\lambda^3_{11} + 0.3\lambda_{11} + 1 - 0.003\lambda^3_{12} + 0.8\lambda_{12} + 3 \\ & - 0.002\lambda^3_{21} + 0.5\lambda_{21} + 2 - 0.001\lambda^3_{22} + 0.3\lambda_{22} + 1 \\ & - 0.001\lambda^3_{31} + 0.3\lambda_{31} + 1 - 0.002\lambda^3_{32} + 0.5\lambda_{32} + 2 \\ & + 0.02\lambda_{11} + 0.04\lambda_{12} + 0.03\lambda_{21} + 0.02\lambda_{22} + 0.02\lambda_{31} + 0.003\lambda_{32} \\ & + (0)(\lambda_4 + \lambda_5) + Z(\lambda_1 + \lambda_2 + \lambda_3) \end{aligned}$$

Expanding the functional gives

$$\begin{aligned} C = {} & 2.32\lambda_{11} + 2.53\lambda_{21} + 2.32\lambda_{31} + 2.34\lambda_{12} + 1.82\lambda_{22} \\ & + 2.03\lambda_{32} + 0.004\lambda_{11}^2 + 0.003\lambda_{21}^2 + 0.004\lambda_{31}^2 + 0.002\lambda_{12}^2 \\ & + 0.004\lambda_{22}^2 + 0.003\lambda_{32}^2 + 0.01\lambda_{11}\lambda_{21} + 0.01\lambda_{11}\lambda_{31} \\ & + 0.01\lambda_{21}\lambda_{31} + 0.01\lambda_{12}\lambda_{22} + 0.01\lambda_{12}\lambda_{32} + 0.01\lambda_{22}\lambda_{32} + 70 \\ & + Z\lambda_1 + Z\lambda_2 + Z\lambda_3 \end{aligned}$$

Table 1 gives constants to be used in the programming calculations.

The data in Table 1 are used in the calculating schedules based on the derivations in the Appendix. Only Tableaux 1, 2, and 13 are shown for brevity. The most economic allocation of manufacturing facilities as determined in the accompanying Tableaux is shown herewith.

CALCULATING SCHEDULE TABLEAU 1

			d_k																	
													-20	-1	-40	-3	-20	-2	-40	-2
			d_i																-1	-40
			b_i										-.004	-.002	-.003	-.003	-.004	-.004	-.004	-.003
													-2.32	-2.34	-2.55	-2.55	-1.82	-2.32	-2.32	-2.03
			a_i	-Z	-Z	-Z	-Z													
		i ↓																		
k →		x_{ko}		λ_1	λ_2	λ_3	λ_4	λ_5	λ_{11}	λ_{12}	λ_{21}	λ_{22}	λ_{31}	λ_{32}						
$= (a_k + \sum_{i=1}^{m} b_i x_{ko})$																				
-Z	λ_1	50		1					1	1										
-Z	λ_2	60			1						1	1								
-Z	λ_3	70				1							1							
0	λ_4	100					1		1		1			1						
0	λ_5	100						1		1		1		1						
1	$C_{Total} = \sum_{k=1}^{m} x_{ko}(a_k + x_{ko} b_k) + d_k$ + Previous $(C_{Total} + \Delta C) = 1802$																			
2	$M = a_n - \sum_{k=1}^{m} x_{kn}(a_k + 2b_k x_{ko}) + 2b_n \lambda_{no} + \sum_{j=1}^{m} \lambda_{jo}(-C_{kj}) x_{kn} + \sum_{k=1}^{m-1} \sum_{j=k+1}^{m} C_{kj} x_{kn} x_{jn}$								-2.32 +Z	-2.34 +Z	-2.53 +Z	-1.82 +Z	-2.32 +Z	-2.03 +Z						
3	$N = b_n + \sum_{k=1}^{m} b_k x_{kn}^2 + \sum_{k=1}^{m} (-C_{kn}) x_{kn} + \sum_{k=1}^{m-1} \sum_{j=k+1}^{m} C_{kj} x_{kn} x_{jn}$								-.004 +Z	-.002 +Z	-.003 +Z	-.004 +Z	-.004 +Z	-.003 +Z						
	Smallest Positive x_{ko}/x_{kn}								50	50	60	60	70	70						
	Optimum $\lambda_n = -M/2N$								-290 +1252	-585 +2302	-421.7 +1662	227.5 +1252	-290 +1252	338.33 +1662						
4	A (Choose according to Table I - Appendix II)								50	50	60	60	70	70						
5	A · N								-.2	-.1	-.18	-.24	-.28	-.21						
6	N + A·N								-2.52 +Z	-2.44 +Z	-2.71 +Z	-2.06 +Z	-2.60 +Z	-2.24 +Z						
7	$\Delta C = A[\sum + A \cdot \overline{N}] + d_n - d_k$ Subtract d_k only if $A = x_{ko}/x_{kn}$ Add d_n only on "unprimed" λ's								-147 +502	-165 +502	-184.6 +602	-164.6 +602	-203 +702	-198.6 +702						

Steps to Follow To Calculate ΔC for λ_n

$$\sum_{k=1}^{m} C_{kn} x_{ko} + \sum_{k=1}^{k=m-1} \sum_{j=k+1}^{m} (-C_{kj})(x_{ko} x_{jn} + x_{kn} x_{jo}) - \lambda_{no} \sum_{k=1}^{m} C_{kn} x_{kn}$$

$$\sum_{k=1}^{k=m-1} \sum_{j=k+1}^{m} C_{kj} x_{jo} x_{ko}$$

CALCULATING SCHEDULE TABLEAU 2

			d_k						—	—	—	—	-42		
			d_i						-20/-1	-40/-3	-20/-2	-40/-1	-20/-1	-40/-2	
			b_i						-.004	-.002	-.003	-.004	-.004	-.003	
$\Pi(g_k + 2b_k x_{ko})$			a_i	-Z	-Z	-Z			-2.32	-2.34	-2.53	-1.82	-2.32	-2.03	
		i →													
	k	x_{ko}		λ_1	λ_2	λ_3	λ_4	λ_5	λ_{11}	λ_{12}	λ_{21}	λ_{22}	λ_{31}	λ_{32}	
-2.45	λ_{32}	70				1							1	1	
-Z	λ_1	50		1					1	1					
-Z	λ_2	60			1						1	1			
0	λ_4	100					1		1		1		1		
0	λ_5	30				-1		1		1		1	-1		

$$\sum_{k=1}^{m} C_{kn} x_{ko} + \sum_{k=1}^{k=m-1} \sum_{j=k+1}^{j=m} (-C_{kj})(x_{ko} x_{jn} + x_{jo} x_{kn}) - \lambda_{no} \sum_{k=1}^{m} C_{kn} x_{kn}$$

Steps to Follow To Calculate ΔC for λ_n

$$\sum_{k=1}^{k=m-1} \sum_{j=k+1}^{j=m} C_{kj} x_{jo} x_{ko}$$

1	$C_{Total} = \sum_{k=1}^{m} \{x_{ko}(a_k + x_{ko} b_k) + d_k\} +$ Previous $(C_{Total} + \Delta C)$	-198.8 -110Z															
2	$M = a_n - \sum_{k=1}^{m} x_{kn}(a_k + 2_{bk} x_{ko}) + 2b_n \lambda_{no} +$					2.45 -Z			-2.32 +Z	-1.64 +Z	-2.53 +Z	-1.12 +Z	+.13				
3	$N = b_n + \sum_{k=1}^{m} b_k x_{kn}^2 + \sum_{k=1}^{m} (-C_{kn}) x_{kn} + \sum_{k=1}^{k=m-1} \sum_{j=k+1}^{j=m} C_{kj} x_{kn} x_{jn}$					-.003			-.004	-.002	-.003	-.004	-.004				
	Smallest Positive x_{ko}/x_{kn}					70			50	30	60	30	70				
	Optimum $\lambda_n = -M/2N$					408.3 -1667Z			-290 +125Z	-410 +250Z	-421.6 +166.7Z	-140 +125Z	9.29				
4	A (Choose according to Table I - Appendix II)					70			50	30	60	30	9.29				
5	$A \cdot N$					-.21			-.2	-.06	-.18	-.12	-.065				
6	$M + A \cdot N$					2.24 -Z			-2.52 +Z	-1.70 +Z	-2.71 +Z	-1.24 +Z	.065				
7	$\Delta C = A[M + A \cdot N] + d_n - d_k$ } Subtract d_k only if $A = x_{ko}/x_{kn}$ Add d_n only on "unprimed" λ's					198.8 -70Z			-147 +50Z	-54.0 +30Z	-184.6 +60Z	-38.2 +30Z	-20.4				

CALCULATING SCHEDULE TABLEAU 13

			d_k	-21	-22	-22		–	–	–	–	–	–	–	–		
			d_j						-20 / -1	-40 / -2	-20 / -2	-40 / -1	-20 / -1	-40 / -2			
			b_j						-.004	-.002	-.003	-.004	-.004	-.003	-.004		
			a_j	-2	-2	-2			-2.32	-2.34	-2.65	-1.82	-2.32	-2.03	-2.32		
$= (a_k + 2b_k x_{ko})$			i →														
	k	x_{ko}		λ_1	λ_2	λ_3	λ_4	λ_5	λ_{11}	λ_{12}	λ_{21}	λ_{22}	λ_{31}	λ_{32}	λ_{31}'		
-2.4183	λ_{12}	19.5712								1							
0	λ_4	20		-1	-1	-1	1	1									
-2.5696	λ_{31}	31.2279											1		-1		
-2.1535	λ_{22}	41.6567				-1		1				1			-1		
-2.3634	λ_{11}	30.4286		1					1								
-2.5401	λ_{21}	18.3433			1	1		-1			1				1		
-2.2626	λ_{32}	38.7721				1								1	1		
1	$C_{Total} = \sum_{k=1}^{m} \left[x_{ko}(a_k + x_{ko} b_k) + d_k \right] +$ Previous $(C_{Total} + \Delta C)$	-427.2955															
2	$M = a_n - \sum_{k=1}^{m} x_{kn} (a_k + 2_{bk} x_{ko}) + 2b_n \lambda_{no} +$			2.07 -2	2.03 -2	2.11 -2		-.4636							.0218		
3	$N = b_n + \sum_{k=1}^{m} b_k x_{kn}^2 + \sum_{k=1}^{m} (-C_{kn}) x_{kn} + \sum_{k=1}^{k=m-1} \sum_{j=k+1}^{j=m} C_{kj} x_{kn} x_{jn}$			-.021	-.004	-.003		-.007							-.034		
	Smallest Positive x_{ko}/x_{kn}			30.43	18.34	18.34		20							18.3433		
	Optimum $\lambda_n = -M/2N$			-23.82	-1892	-166.72		-32.4							.3206		
4	A (Choose according to Table I - Appendix II)			30.43	18.34	18.34		20							.3206		
5	$A \cdot N$			-.64	-.07	-.06		-.14							-.0109		
6	$M + A \cdot N$			1.43 -2	1.96 -2	2.05 -2		-.5936							.0109		
7	$\Delta C = A[M + A \cdot N] + d_n - d_k$ Subtract d_k only if $A = x_{ko}/x_{kn}$ Add d_n only on "unprimed" λ's			44.81 -30.432	37.96 -18.342	39.60 -18.342		-11.9							.0039		

$$\sum_{k=1}^{m} C_{kn} x_{ko} + \sum_{k=1}^{k=m-1} \sum_{j=k+1}^{j=m} (-C_{kj})(x_{ko} x_{jn} + x_{jo} x_{kn}) - \lambda_{no} \sum_{k=1}^{m} C_{kn} x_{kn}^2$$

$$\sum_{k=1}^{k=m-1} \sum_{j=k+1}^{j=m} C_{kj} x_{jo} x_{ko}$$

Steps to Follow To Calculate ΔC for λ_n

TABLE 1 CONSTANTS FOR PROGRAMMING CALCULATIONS

Variable	a	b	c	d
λ_{11}	2.32	0.004	..	1 + 20 if plant 1 is in system
λ_{21}	2.53	0.003	..	2 + 20 if plant 1 is in system
λ_{31}	2.32	0.004	..	1 + 20 if plant 1 is in system
λ_{12}	2.34	0.002	..	3 + 40 if plant 2 is in system
λ_{22}	1.82	0.004	..	1 + 40 if plant 2 is in system
λ_{32}	2.03	0.003	..	2 + 40 if plant 2 is in system
$\lambda_{11}\lambda_{21}$	..	...	0.01	.....
$\lambda_{11}\lambda_{31}$	..	...	0.01	.....
$\lambda_{21}\lambda_{31}$	..	...	0.01	.....
$\lambda_{12}\lambda_{22}$	..	...	0.01	.....
$\lambda_{12}\lambda_{32}$	..	...	0.01	.....
$\lambda_{22}\lambda_{32}$	..	...	0.01	.....

Consuming points \ Manufacturing plants	1	2	Required amount specified at each consuming point
1	30.43	19.57	50
2	18.34	41.66	60
3	31.23	38.77	70
Total at each manufacturing plant	80.00	100.00	180

Total minimum cost = 427.89

Appendix 2

Testing the Optimality of the Solution

To test the optimality of the solution as given by the vectors ($\lambda_1 - \lambda_m$) in the basis, let λ_n be included in the basis so that the constraining Equations [2] become

$$\left.\begin{aligned} x_{10} &= \lambda_1 + x_{1n}\lambda_n \\ x_{20} &= \lambda_2 + x_{2n}\lambda_n \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{k0} &= \lambda_k + x_{kn}\lambda_n \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{m0} &= \lambda_m + x_{mn}\lambda_n \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots [3]$$

The functional Equation [1] then becomes

$$C = \sum_{i=1}^{n} a_i\lambda_i + \sum_{i=1}^{n} b_i\lambda_i^2 + \sum_{\substack{i=1 \\ =i+1}}^{\substack{j=n \\ i=n-1}} c_{ij}\lambda_i\lambda_j + \sum_{i=1}^{n} \frac{d_i\lambda_i}{\lambda_i + \delta} \quad \ldots\ldots [4]$$

The rate of change of the functional (C) with λ_n is

$$\frac{dC}{d\lambda_n} = \sum_{i=1}^{n} a_i \frac{d\lambda_i}{d\lambda_n} + \sum_{i=1}^{n} 2b_i\lambda_i \frac{d\lambda_i}{d\lambda_n} + \sum_{\substack{i=1 \\ j=i+1}}^{\substack{j=n \\ i=n-1}} c_{ij}\left[\lambda_i \frac{d\lambda_j}{d\lambda_n} + \lambda_j \frac{d\lambda_i}{d\lambda_n}\right] + \sum_{i=1}^{n} \frac{(\lambda_i + \delta)d_i \frac{d\lambda_i}{d\lambda_n} - d_i\lambda_i \frac{d\lambda_i}{d\lambda_n}}{(\lambda_i + \delta)^2} \quad \ldots\ldots [5]$$

From the constraining Equations [3]

$$\lambda_k = x_{k0} - x_{kn}\lambda_n \quad \text{and} \quad \frac{d\lambda_k}{d\lambda_n} = -x_{kn}$$

Substituting these values into Equation [5] and noting that $d\lambda_i/d\lambda_n = 1$ for $i = n$, the rate of change of C with λ_n becomes

$$\frac{dC}{d\lambda_n} = a_n + \sum_{i=1}^{m}\left\{(-x_{in})(a_i + 2b_ix_{i0}) + c_{in}x_{i0}\right\} + \frac{\delta d_n}{(\lambda_n + \delta)^2} + \sum_{\substack{i=1 \\ j=i+1}}^{\substack{j=m \\ i=m-1}} (-c_{ij})(x_{i0}x_{jn} + x_{j0}x_{in}) + 2\lambda_n\left\{b_n + \sum_{i=1}^{m}(b_ix_{in}^2 - c_{in}x_{in}) + \sum_{\substack{i=1 \\ j=i+1}}^{\substack{j=m \\ i=m-1}} c_{ij}x_{in}x_{jn}\right\} + \sum_{i=1}^{m} \frac{\delta d_i(-x_{in})}{(x_{i0} - x_{in}\lambda_n + \delta)^2} \quad \ldots\ldots [6]$$

The increment ΔC_n of the functional is the total change in the functional due to substituting λ_n into the basis and adjusting the other variables in the basis to satisfy the constraining equations. The increment ΔC_n of the functional is simply the definite integral

$$\Delta C_n = \int_0^{\lambda_n} \frac{dC}{d\lambda_n}\, d\lambda_n \quad \ldots\ldots [7]$$

The limits 0 and λ_n follow from the realization that initially λ_n is not in the basis and hence is equal to zero; and that the final value of the variable λ_n is a definite value (which we call λ_n) to be determined by methods described later. If C_0 is the value of the functional with $\lambda_n = 0$, then with λ_n in the basis, the new value of the functional is

$$C_n = C_0 + \Delta C_n$$

Substitute λ_n into the basis and let the final value of $\lambda_n = x_{k0}/x_{kn}$ (to be explained subsequently). Then the increment of the functional ΔC_n is found by evaluating the definite integral, Equation [7], and letting $\delta \to 0$

$$\Delta C_n = \int_0^{x_{k0}/x_{kn}} \frac{dC}{d\lambda_n}\, d\lambda_n = \frac{x_{k0}}{x_{kn}}\left[a_n + \sum_{i=1}^{m}\left\{(-x_{in})(a_i + 2b_ix_{i0}) + c_{in}x_{i0}\right\} + \sum_{\substack{i=1 \\ j=i+1}}^{\substack{j=m \\ i=m-1}} (-c_{ij})(x_{i0}x_{jn} + x_{j0}x_{in}) + \frac{x_{k0}}{x_{kn}}\left\{b_n + \sum_{i=1}^{m}(b_ix_{in}^2 - c_{in}x_{in}) + \sum_{\substack{i=1 \\ j=i+1}}^{\substack{j=m \\ i=m-1}} c_{ij}x_{in}x_{jn}\right\}\right] + d_n - d_k \quad \ldots\ldots [7a]$$

Equation [7a] is of the form

$$\Delta C_n = \frac{x_{k0}}{x_{kn}}\left[M + \frac{x_{k0}}{x_{kn}}N\right] + P$$

M is simply the rate of change of the functional with λ_n at $\lambda_n = 0^+$ (since the functional has a finite step in value at $\lambda_n = 0$, the rate of change at $\lambda_n = 0$ is infinite). M can be evaluated by letting $\lambda_n = 0^+$ and $\delta = 0$ in Equation [6]. N is an indirect contribution to the increment of the functional. P is the difference between the constant terms of the functional associated with λ_n and λ_k. The values of M, N, and P are independent of x_{k0}/x_{kn} and depend only on the values of the constants a, b, c, and d and the values of the other variables in the basis. This nomenclature will be utilized on the calculating schedules.

The existence of the derivative of the functional with respect to λ_n implies the possibility of a maximum or minimum value for the functional. The optimum value of λ_n (the amount of λ_n that makes C a maximum or minimum) is found from Equation [6] by solving $(dC)/(d\lambda_n) = 0$ with $\delta = 0$. Hence

$$\text{Optimum } \lambda_n = \frac{-a_n - \sum_{i=1}^{m}\left\{(-x_{in})(a_i + 2b_i x_{i0}) + c_{in}x_{i0}\right\} - \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m}(-c_{ij})(x_{i0}x_{jn} + x_{j0}x_{in})}{2\left\{b_n + \sum_{i=1}^{m}(b_i x_{in}^2 - c_{in}x_{in}) + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m} c_{ij}x_{in}x_{jn}\right\}} = \frac{-M}{2N} \quad \ldots [6a]$$

The usefulness of the optimum λ_n depends on its numerical value relative to x_{k0}/x_{kn}. The values of λ_n must satisfy the inequality $0 \leq \lambda_n \leq x_{k0}/x_{kn}$, where x_{k0}/x_{kn} is the minimum positive value for $k = 1, \ldots, m$ (as shown subsequently this choice of x_{k0}/x_{kn} assures that the variables $\lambda_1, \ldots, \lambda_m$ will be zero or positive). The value of λ_n to be used in forming the new basis is the smaller positive value of either the optimum λ_n or x_{k0}/x_{kn}.

Using λ_n to Improve the Solution if $x_{kn}/x_{k0} \leq$ Optimum λ_n

The method of finding the λ_n to improve the solution is to calculate ΔC_i for each λ_i, $i = n, n+1, \ldots, r$ and to choose as λ_n the λ_i which gives the largest ΔC_i. Since the basis can have only m-vectors, λ_n must be substituted for a vector in the basis. The vector to be removed from the basis is λ_k which corresponds to x_{k0}/x_{kn} (minimum positive x_{k0}/x_{kn} for $k = 1, 2, \ldots, m$).

The substitution of λ_n for λ_k in the basis is made according to the method of elimination. Consider the equation

$$x_{k0} = \lambda_k + x_{kn}\lambda_n + x_{k(n+1)}\lambda_{(n+1)} + \ldots + x_{kr}\lambda_r \ldots . [8]$$

Divide by x_{kn}, the coefficient of λ_n. Then

$$\frac{x_{k0}}{x_{kn}} = \frac{\lambda_k}{x_{kn}} + \lambda_n + \frac{x_{k(n+1)}}{x_{kn}}\lambda_{(n+1)} + \ldots + \frac{x_{kr}}{x_{kn}}\lambda_r \ldots [8a]$$

Now $\lambda_k = \lambda_{n+1} = \lambda_{n+2} = \ldots = \lambda_r = 0$, since they are not in the basis. Hence the value of λ_n in the basis is x_{k0}/x_{kn}.

Since λ_n has a particular value in the kth equation, the remaining $(m - 1)$ equations can be adjusted so that λ_n does not appear in them. That is, to be consistent with the form of Equations [2], λ_n should appear in only one equation in the set. This adjustment is made very simply by the elimination method. For example, the jth equation is

$$x_{j0} = \lambda_j + x_{jn}\lambda_n + x_{j(n+1)}\lambda_{n+1} + \ldots + x_{jr}\lambda_r \ldots . [8b]$$

Multiply Equation [8a] by x_{jn} and subtract the resulting equation from Equation [8b] to obtain

$$x_{j0} - x_{jn}\frac{x_{k0}}{x_{kn}} = \lambda_j - \frac{x_{jn}}{x_{kn}}\lambda_k + \left(x_{j(n+1)} - \frac{x_{jn}x_{k(n+1)}}{x_{kn}}\right)\lambda_{n+1} + \ldots + \left(x_{jr} - x_{jn}\frac{x_{kr}}{x_{kn}}\right)\lambda_r \ldots \ldots [8c]$$

Equation [8a] and Equation [8c] constitute a new set of equations similar to Equations [2]. The basis of the new set of equations is an improved solution and gives an improved value for C.

Examination of the constant term of Equation [8c] shows why x_{k0}/x_{kn} must be the minimum positive value. Rewrite the constant term of Equation [8c] as

$$x_{jn}\left[\frac{x_{j0}}{x_{jn}} - \frac{x_{k0}}{x_{kn}}\right]$$

Since $x_{j0}/x_{jn} \geq x_{k0}/x_{kn}$, the constant term is positive and $\lambda_j \geq 0$. This is an assumed restriction as stated in the definition of the problem.

New values of ΔC are calculated for the equations with the new basis and the solution is improved by substituting a new λ_n into the basis. The calculations are repeated until ΔC is no longer favorable, i.e., $\Delta C \geq 0$.

Using λ_n to Improve the Solution if Optimum $\lambda_n \leq x_{k0}/x_{kn}$

If optimum $\lambda_n \leq x_{k0}/x_{kn}$, a supplementary equation should be added since another variable λ_n is entering the basis but no variable is being removed from the basis. A variable is not removed from the basis if $\lambda_n < x_{k0}/x_{kn}$ since (see Equation [8a]) in that case

$$\lambda_k = x_{kn}\left(\frac{x_{k0}}{x_{kn}} - \lambda_n\right)$$

The supplementary equation is a transformation

$$\lambda_n = \lambda_{n0} + \lambda_n'$$

where λ_n is the original variable, λ_{n0} is the optimum value of λ_n to be added into the basis, and λ_n' is a new variable. The original m-equations are modified by substituting the transformation in each of the original m-equations, [8b], in order to eliminate λ_n and to introduce λ_n'. The new set of $(m + 1)$ equations consists of the supplementary equation and the m-equations from which λ_n has been eliminated and to which λ_n' has been added. The constants associated with λ_n' are a_n, b_n, and c_{in}. The variable λ_n' can be either positive or negative as will be shown subsequently.

The method of continuing the calculations after λ_n' has been introduced into the problem is similar in many respects to the routine calculation of ΔC previously described. For all of the variables $\lambda_{n+1}, \ldots, \lambda_r$ calculate ΔC as described previously using the original m-equations. To calculate ΔC for λ_n' a technique derived as follows may be used:

From Equation [1]

$$C = a_n\lambda_n + b_n\lambda_n^2 + \sum_{i=1}^{m} c_{in}\lambda_i\lambda_n + \frac{d_n\lambda_n}{\lambda_n + \delta} + \sum_{i=1}^{m} a_i\lambda_i + \sum_{i=1}^{m} b_i\lambda_i^2 + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m} c_{ij}\lambda_i\lambda_j + \sum_{i=1}^{m}\frac{d_i\lambda_i}{\lambda_i + \delta}$$

Let $\lambda_n = \lambda_{n0} + \lambda_n'$ where λ_{n0} = optimum λ_n previously calculated. Then

$$C = a_n(\lambda_{n0} + \lambda_n') + b_n(\lambda_{n0} + \lambda_n')^2 + \sum_{i=1}^{m} c_{in}\lambda_i(\lambda_{n0} + \lambda_n') + \frac{d_n(\lambda_{n0} + \lambda_n')}{(\lambda_{n0} + \lambda_n') + \delta} + \sum_{i=1}^{m} a_i\lambda_i + \sum_{i=1}^{m} b_i\lambda_i^2 + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m} c_{ij}\lambda_i\lambda_j + \sum_{i=1}^{m}\frac{d_i\lambda_i}{\lambda_i + \delta} \quad \ldots [9]$$

The rate of change of C with λ_n' is

$$\frac{dC}{d\lambda_n'} = a_n + 2b_n(\lambda_{n0} + \lambda_n') + \sum_{i=1}^{m} c_{in}\lambda_{n0}\frac{d\lambda_i}{d\lambda_n'} + \sum_{i=1}^{m} c_{in}\lambda_i + \sum_{i=1}^{m} c_{in}\lambda_n'\frac{d\lambda_i}{d\lambda_n'} + \frac{(\lambda_{n0} + \lambda_n' + \delta)d_n - d_n(\lambda_{n0} + \lambda_n')}{(\lambda_{n0} + \lambda_n' + \delta)^2} + \sum_{i=1}^{m} a_i\frac{d\lambda_i}{d\lambda_n'} + \sum_{i=1}^{m} 2b_i\lambda_i\frac{d\lambda_i}{d\lambda_n'} + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m} c_{ij}\left[\lambda_i\frac{d\lambda_j}{d\lambda_n'} + \lambda_j\frac{d\lambda_i}{d\lambda_n'}\right] + \sum_{i=1}^{m}\frac{d_i(\lambda_i + \delta)\frac{d\lambda_i}{d\lambda_n'} - d_i\lambda_i\frac{d\lambda_i}{d\lambda_n'}}{(\lambda_i + \delta)^2} \quad \ldots\ldots [10]$$

From the m restricting equations (similar to Equation [3] with λ_n' in place of λ_n)

$$\lambda_i = x_{i0} - x_{in}'\lambda_n', \text{ and } \frac{d\lambda_i}{d\lambda_n'} = -x_{in}'$$

Substitute for λ_i and $(d\lambda_i)/(d\lambda_n')$ in Equation [10] and rearrange to obtain

$$\text{Optimum } \lambda_n' = \frac{-a_n - \sum_{i=1}^{m}\left[(-x_{in}')(a_i + 2b_ix_{i0}) + c_{in}x_{i0}\right] - \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m}(-c_{ij})(x_{i0}x_{jn}' + x_{j0}x_{in}') - 2b_n\lambda_{n0} + \sum_{i=1}^{m} c_{in}x_{in}'\lambda_{n0}}{2b_n + 2\sum_{i=1}^{m}(b_ix_{in}'^2 - c_{in}x_{in}') + 2\sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m} c_{ij}x_{jn}'x_{in}'} \approx \frac{-M'}{2N} \quad \ldots\ldots [12]$$

$$\frac{dC}{d\lambda_n'} = a_n + \sum_{i=1}^{m}\left[(-x_{in}')(a_i + 2b_ix_{i0}) + c_{in}x_{i0}\right] + \frac{\delta d_n}{(\lambda_{n0} + \lambda_n' + \delta)^2} + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m}(-c_{ij})(x_{i0}x_{jn}' + x_{j0}x_{in}') + 2\lambda_n'\left[b_n + \sum_{i=1}^{m}(b_ix_{in}'^2 - c_{in}x_{in}') + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m} c_{ij}x_{jn}'x_{in}'\right] + \sum_{i=1}^{m}\frac{-\delta d_ix_{in}'}{(x_{i0} - x_{in}'\lambda_n' + \delta)^2} + 2b_n\lambda_{n0} - \sum_{i=1}^{m} c_{in}x_{in}'\lambda_{n0} \quad [10a]$$

The rate of change of the functional for $\lambda_n' = 0$ is found from Equation [10a] by letting $\lambda_n' = 0$ and $\delta = 0$. Then

$$\left[\frac{dC}{d\lambda_n'}\right]_{\lambda_n'=0} = a_n + \sum_{i=1}^{m}\left[(-x_{in}')(a_i + 2b_ix_{i0}) + c_{in}x_{i0}\right] + \sum_{i=1}^{i=m-1}\sum_{j=i+1}^{j=m}(-c_{ij})(x_{i0}x_{jn}' + x_{j0}x_{in}') + 2b_n\lambda_{n0} - \sum_{i=1}^{m} c_{in}x_{in}'\lambda_{n0} = M' \quad \ldots\ldots [10b]$$

Consider Equation [10a] as divided into two parts: Part 1 is the first six terms and Part 2 is the last two terms. Part 1 is similar to Equation [6] with λ_n' substituted for λ_n except for the term

$$\frac{\delta d_n}{(\lambda_{n0} + \lambda_n' + \delta)^2}$$

Hence, integrating Part 1 with respect to λ_n' between the limits of 0 and λ_n', substituting $\lambda_n' = x_{i0}/x_{in}'$ (as found using the x's in the set of $m + 1$ equations with λ_n') and letting $\delta \to 0$, we obtain Equation [7a] with λ_n' substituted for λ_n and without the term d_n (d_n is essentially a constant term already included with λ_n in the total value of the functional).

Integrating Part 2 between the same limits gives

$$2b_n\lambda_{n0}\lambda_n' - \sum_{i=1}^{m} c_{in}x_{in}'\lambda_{n0}\lambda_n'$$

Therefore the total incremental gain found by substituting $\lambda_n' = x_{i0}/x_{in}'$ is

$$\Delta C = (\text{Equation [7}a\text{] without } d_n) + 2b_n\lambda_{n0}\left(\frac{x_{i0}}{x_{in}'}\right) - \sum_{i=1}^{m} c_{in}x_{in}'\lambda_{n0}\frac{x_{i0}}{x_{in}'} \quad \ldots\ldots [11]$$

The form of Equation [11] is the same as Equation [7a] so that the calculation schedule is arranged for both calculations.

The optimum amount of λ_n' is found by solving $(dC)/(d\lambda_n') = 0$ for λ_n'. This calculation yields

The amount of the variable λ_n' (λ_n' can be either positive or negative) to substitute into the basis depends on the relative values of the optimum λ_n', x_{i0}/x_{in}', and

$$M = \left[\frac{dC}{d\lambda_n'}\right]_{\lambda_n'=0}$$

These three quantities describe the geometry of the functional for various values of λ_n'.

Table 2 (based on the geometry of C) indicates the value of λ_n or λ_n' to substitute into the basis and indicates the value to be used in Step 4 of the calculating schedule.

The method of handling optimum quantities can be generalized to include any amount of a variable less than x_{i0}/x_{in}. The need for this generalization arises in problems where one or more of the variables are restricted as follows

$$\lambda_i = 0 \text{ or } g \leq \lambda_i \leq h$$

where g and h are positive numbers.

Arranging the Linear Restricting Equations

In most cases the constraining equations are not initially in the form of Equation [2]. They are placed in this form to facilitate the solution. Assume the equations describing the the system are

$$\left.\begin{aligned} x_{10} &= x_{1n}\lambda_n + \ldots + x_{1r}\lambda_r \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{k0} &= x_{kn}\lambda_n + \ldots + x_{kr}\lambda_r \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{m0} &= x_{mn}\lambda_n + \ldots + x_{mr}\lambda_r \end{aligned}\right\} \quad \ldots\ldots\ldots [13]$$

TABLE 2 SELECTION OF VALUE FOR λ_n OR λ_n' FOR CALCULATION OF ΔC^*

Case no.	M (Eq. [10b])	Optimum value (Equation [12])		Value of λ_n or λ_n' (Value of A in Step 4 of calculating schedule)
		Sign	Value	
1	+	+	$< \frac{x_{k0}}{x_{kn}}$	Optimum λ_n
2	+	+	$\geq \frac{x_{k0}}{x_{kn}}$	$\frac{x_{k0}}{x_{kn}}$
3	+	−	$\left\lvert\frac{\text{Opt}}{\lambda_n}\right\rvert < \lambda_{n0}$	Try both $-\lambda_{n0}$ and $\frac{x_{k0}}{x_{kn}}$
4	+	−	$\left\lvert\frac{\text{Opt}}{\lambda_n}\right\rvert \geq \lambda_{n0}$	$\frac{x_{k0}}{x_{kn}}$
5	−	+	$\left\lvert\frac{\text{Opt}}{\lambda_n}\right\rvert < \frac{x_{k0}}{x_{kn}}$	Try both $-\lambda_{n0}$ and $\frac{x_{k0}}{x_{kn}}$
6	−	+	$\left\lvert\frac{\text{Opt}}{\lambda_n}\right\rvert \geq \frac{x_{k0}}{x_{kn}}$	$-\lambda_{n0}$
7	−	−	$\left\lvert\frac{\text{Opt}}{\lambda_n}\right\rvert < \lambda_{n0}$	Optimum λ_n
8	−	−	$\left\lvert\frac{\text{Opt}}{\lambda_n}\right\rvert \geq \lambda_{n0}$	$-\lambda_{n0}$

* $\lambda_{n0} \geq 0$

Equations [13] can be transformed into Equations [2] by adding m variables $\lambda_1 \ldots \lambda_m$. The significance of $\lambda_1 \ldots \lambda_m$ is that they are artificial vectors which must be zero for the optimum performance. Hence in the functional they are assigned a cost of (M) or a profit $(-M)$, where M is a number much, much larger (unbounded) than any other number in the calculation. Hence it will be unprofitable to have the artificial vectors in the optimum solution.

Assume that the initial restrictions are inequations such as

$$\left.\begin{aligned} x_{10} &> x_{1n}\lambda_n + \ldots + x_{1r}\lambda_r \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{k0} &> x_{kn}\lambda_n + \ldots + x_{kr}\lambda_r \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{m0} &> x_{mn}\lambda_n + \ldots + x_{mr}\lambda_r \end{aligned}\right\} \quad \ldots\ldots\ldots [14]$$

These inequations can be transformed to equations by adding m "slack" vectors $\lambda_1 \ldots \lambda_m$ to give

$$\left.\begin{aligned} x_{10} &= \lambda_1 + x_{1n}\lambda_n + \ldots + x_{1r}\lambda_r \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ &\text{—} \\ x_{m0} &= \lambda_m + x_{mn}\lambda_n + \ldots + x_{mr}\lambda_r \end{aligned}\right\} \quad \ldots\ldots [14a]$$

These equations are of the required form of Equations [2]. The slack vectors absorb the excess of x_{k0} which are not required by $\lambda_n \ldots \lambda_r$. Since the slack vectors do not affect costs or profits, they enter the functional with zero cost or zero profit.

Some Observations on Formal Models for Programming

By J. J. SLADE, JR.,[1] NEW BRUNSWICK, N. J.

The problems of linear and nonlinear programming are reviewed using a geometrical representation, with a general discussion of the various types of solutions arising to the problem of finding a maximum or minimum of a property $z = f(x_1, \ldots, x_n)$ when the variables x_i are subject to a certain number of constraints. A special type of nonlinear programming is considered when the departure from linearity is due to a systematic drift toward a saturation level as is the case when profit actually decreases after the rate of production attains its optimum value. The mathematical model satisfying this requirement is one for which a linear factor dominates the profit function for small rates of production and a negative exponential for high rates of production. A solution with illustrative examples is presented to this problem.

Introduction

A VERY old problem in mathematics is that of finding the values of the components of a multicomponent system at which some designated property of this system attains an optimum value, when the components satisfy certain constraints. Programming, as currently understood, differs from the classical problem only in that the optimum value of the property is sought within a region that satisfies a set of inequalities rather than a fixed locus represented by a set of equations.

A formal statement of the problem of programming is the following: Let $(x_1, x_2, \ldots, x_n)$ designate an n-component complex. It is required to maximize (or minimize) the property

$$z = f(x_1, x_2, \ldots x_n) \quad \ldots\ldots [1]$$

subject to the $n + k$ constraints

$$\left.\begin{array}{c} 0 \leq x_1, \quad 0 \leq x_2, \ldots, 0 \leq x_n \\ g_1(x_1 \ldots x_n) \leq h_1, \ldots, g_k(x_1, \ldots x_n) \leq h_k \end{array}\right\} \quad \ldots\ldots [2]$$

In dealing with problems of this nature certain types of difficulties are encountered which must be recognized explicitly and formulated. Some of these are merely mathematical complexities. Others have deeper sources. They have to do with the initial formulation of the problem; with whether Equation [1] and the Inequalities [2] do indeed present a model for the physical system that confronts the investigator.

Among the purely mathematical questions to be answered are those that have to do with the existence of solutions. It appears that for very general situations these questions have been answered.[2] Others have to do with the physical routine required to obtain numerical answers to these problems. These latter difficulties are principally arithmetical in nature. Successful algorithms must be constructed with reference to the character of the function $f(x_1, \ldots x_n)$ of Equation [1] and the g's of the Inequalities [2]. An algorithm has been elaborately investigated in the case in which both f and the g's are linear.[3] Even in this simplest of formulations, however, the amount of computation required would often be forbidding were it not for the existence of the many computational aids now available.

The Geometrical Model

Although purely mathematical questions are not further considered in this paper, it seems desirable to discuss briefly and in a purely descriptive manner, what is implied in the solution of a programming problem.

For ease of visualization or for economy of expression, the problem of programming is considered a problem in geometry by mathematicians, this in spite of the fact that the components $x_1, \ldots x_n$ of the complex, called a vector, must then be co-ordinates in an n-dimensional space. The property to be optimized represented by the equation $z = f(x_1, \ldots x_n)$, called a functional, is itself an n-dimensional hypersurface in the $(n + 1)$-dimensional space $x_1 \ldots x_n, z$. Unfortunately, those who like to draw pictures of their space configurations must be satisfied with two-dimensional vectors (x_1, x_2), the third dimension of the space available to them being taken up by the ordinate z. In order to explain certain mathematical features of the problem we shall restrict our discussion to a 2-space in which x_1, x_2 will be called x, y. The function $z = f(x, y)$ will then be a surface (a 2-space itself) in the 3-space (x, y, z).

In Fig. 1(A) the linear configuration $z = ax + by + c$ is shown. We consider the $n + k = 2 + 2 = 4$ constraining inequalities $0 \leq x, 0 \leq y, b_{11}x + b_{12}y \leq b_{10}, b_{21}x + b_{22}y \leq b_{20}$. The points that satisfy these inequalities are in the region bounded by the lines (1-spaces)

$$x = 0, \quad y = 0, \quad b_{11}x + b_{12}y = b_{10}, \quad b_{21}x + b_{22}y = b_{20} \ldots [3]$$

This is the shaded region marked C in the x-y plane. The shaded area of the plane $z = ax + by + c$ above C includes all the values of the functional at the points of the constraining set C. A plane can have no maximum. The value of z rises along every line not parallel to the x-y plane, provided the line is traversed in the proper direction. Maxima of z can exist only because there are boundaries which cannot be crossed. It is obvious that a maximum (or a minimum) can occur at the intersections p, q, r, and O of the boundaries. The slope of the plane will determine which of these points is a solution. If, however, one boundary happens to be parallel to the x-y plane, then all its points are at the same level, so that all the points of a line segment may be a solution. Every point of the set C is a solution in the exceptional case when the plane $z = ax + by + c$ is parallel to the x-y plane.

One other situation may arise not depicted in this sketch. Suppose, for example, that there are three constraints $0 \leq x$, $0 \leq y \leq d$ and that the plane is not parallel to the x-axis ($a \neq 0$).

---

[1] Research Director, Bureau of Engineering Research, College of Engineering, Rutgers University.

[2] "Non-Linear Programming," by H. W. Kuhn and A. W. Tucker, Second Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, Calif., 1951, pp. 481–492.

Contributed by the Management Division and presented at a joint session of the Management and Production Engineering Divisions at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 26, 1954. Paper No. 54—A-241.

[3] "An Introduction to Linear Programming," by W. W. Cooper, A. Henderson, and A. Charnes, John Wiley & Sons, Inc., New York, N. Y., 1953.

Depending on the slope of the plane, either a maximum occurs at $x = 0$ or else there is no maximum, the value of z continuing to rise as $x \to \infty$, since x is not constrained. If $a = 0$, then the maximum occurs at $y = 0$ or $y = d$, depending on the slope, with x arbitrary.

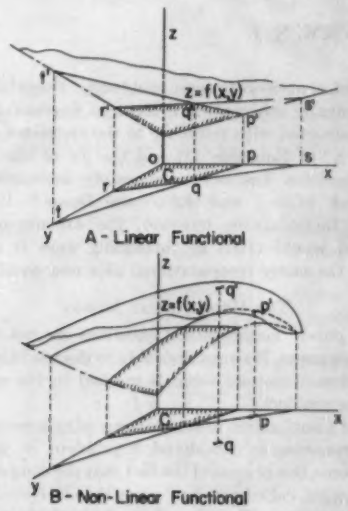


FIG. 1 RELATION OF FUNCTIONAL TO CONSTRAINING REGION

It should be noted that not all the bounding Equations [3] may be satisfied simultaneously. In the case of Fig. 1(A) one pair at a time may be satisfied, yielding the point of intersection of the corresponding lines. In the general case of $n + k$ constraints not more than n-equations may be generally satisfied simultaneously.[4]

The two-dimensional configuration shows all the features that may be expected in the linear problem, regardless of its dimensionality. Each linear equation obtained from an inequality represents an $(n - 1)$-space. The intersection of r of these spaces, represented by r simultaneous equations, is an $(n - r)$-space. If an $(n - r)$-space is "parallel" to the space of the co-ordinates, then all its points may be solutions of the problem. If none is parallel, then the solution is at one of the points of intersection of n of the $n + k$ bounding spaces. There are $(n + k)!/n!k!$ of these points, but not all are acceptable solution. For example, in Fig. 1(A) there are $(2 + 2)!/2!2! = 6$ points of intersection of the bounding 1-spaces, but only four are acceptable, since the two points s and t fall outside the boundary of C.

The general problem could be solved by determining all the acceptable points of intersection, then substituting the sets of values of the co-ordinates in f and picking out that set for which f is greatest (or least). The number of these points soon becomes very large, however, and when it is remembered that the determination of each point involves, in effect, the solution of n simultaneous equations, it will be realized that the labor required to obtain a solution is very great. The methods of solution that have been developed are designed to reduce the number of trial solutions to a minimum.

In Fig. 1(B) is sketched a nonlinear functional of two variables. Its maximum is to be found subject to the same constraints as before. Aside from the fact that an algorithm devised to solve a problem of this nature may be much more laborious than that of the linear case, the essential complication that enters here is that the maximum need not be at a point intersection of the constraining boundaries; in fact, it need not be on a boundary at all. For example, if q, where a maximum of $z = f(x, y)$ occurs, were inside C, it would be a solution of the problem. The determination of this point would involve no statement of the inequalities, other than that required to determine whether q is, or is not, a point of C. A maximum may occur on one of the constraining lines, such as at p. This would be a solution of the problem if p happened to be on the boundary of C. In the general nonlinear case a maximum may occur on any $(n - r)$-dimensional bounding space. The complications arising from this situation can become exceedingly great, for not only must the values of z be tested at the point intersections, but also at the points of the lines, planes, 3-spaces, and so on, determined by all acceptable combinations of the bounding $(n - 1)$-spaces, on the bounding spaces themselves, and on the whole n-space in which maxima may occur independently of the constraints. This discussion indicates that the introduction of nonlinearity into the problem may increase the order of difficulty by manyfold.

The first approximation in the solution of a physical problem is the linear formulation. Mathematically the linear relation is the simplest of relations, and it is often mathematically expedient to attempt a linear formulation of a problem. It is a geometrical fact that an arc may be approximated by a tangent line segment if the distance traversed from the point of tangency is sufficiently small. Analytically this amounts to dropping the square and higher powers in the Taylor expansion of a function. It thus often happens that within the ranges of variation of the variables that have practical significance the functions investigated are practically linear. Sometimes linearity persists over surprisingly long ranges and we have, for example, Hooke's law and Ohm's law.

If we consider the variation of a quantity z due to the variation of a single component x, the variation is said to be linear in both cases A and B of Fig. 2 where M and N are functions of the other components of the system and a is a constant. The two lines in each case represent the variation of z with x for two values of M and N. It is only when the variation due to each component is as in A that z is linear. The surface $z = xy$ is an example of case B.

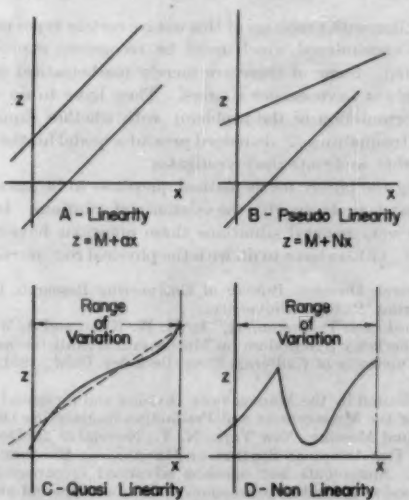


FIG. 2 TYPES OF VARIATION OF THE FUNCTIONAL

[4] A good discussion of simultaneous linear equations and linear transformations is found in "Introduction to Higher Algebra," by M. Bocher, The Macmillan Company, New York, N. Y., 1927, Chapters III-IV. A more geometrical approach is found in "A Survey of Modern Algebra," by G. Birkhoff and A. McLane, The Macmillan Company, New York, N. Y., 1953, Chapter VII, et seq.

The co-ordinates x, z are on a straight line along the lines $y =$ const, and the co-ordinates y, z are on a straight line along $x =$ const, but the surface is not a plane. Along a line $y = mx$, for example, $z = mx^3$ is a parabola.

A useful linear formulation occurs when the variation of a quantity within the important range of variation of the component x does not fluctuate disastrously about a straight line. This quasi-linear condition is shown in Fig. 2(C). The fluctuations shown in D of the same figure are not essentially different from that of C, and a straight line may be found that most nearly fits this curve within the indicated range of variation. The fluctuations about linearity can be said to be large or small only with respect to the vertical origin and scale. If at a point x the value of the functional is z and that of the "best straight line" functional is z', then the fluctuation is quasi-linear if $|z - z'|/z'$ is acceptably small. But even when this ratio is small the function represented in D cannot be approximated by a straight line if features such as the cusp and trough are significant. When features occur that rule out the linear formulation a model must be found that incorporates them while, if possible, retaining a measure of mathematical simplicity.

The Model With Saturation

In particular, a feature frequently and quite universally encountered is the phenomenon of saturation. Often the departure from linearity is due to a systematic drift toward a saturation level that is significant. This feature will be examined in some detail. A model that accounts for saturation in a multicomponent field will be described and its significant properties inferred. No attempt is made to develop an algorithm for the numerical solution of specific problems, but it will be shown that, though the model is nonlinear, only linear processes are required in the solution of the optimizing problem.

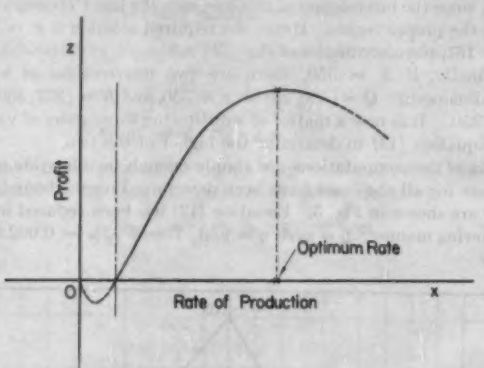


Fig. 3 The Phenomenon of Saturation

A curve that represents saturation in a one-component system is sketched in Fig. 3. To give the depicted variation some semblance of physical substance the component will be called the rate of production of some item and the functional will be called the profit. When the rate of production and profit are linearly connected the profit increases proportionately with the rate without limit. In the figure the two quantities are linearly related near the second zero crossing of the curve, but the deviation from linearity is persistent as higher rates are considered; that is, in a system such as this a high rate of production may be maintained only by sacrificing profit.

An outside constraint may be introduced that limits the rate of production. Such would be the inherent capacity of the producing unit. If this constraint is severe enough the effect of saturation might be negligible at the highest permissible rates of production and the problem would be essentially linear. Without such an outside constraint, or with one that permits high rates of production, saturation must certainly be taken into account. In fact, the rate at which maximum profit is obtained is a natural internal constraint of the system. This sketch indicates the sort of generalization we wish to make. For simplicity we may either neglect to incorporate in the generalization the loss that may occur at extremely low rates of production, or we may take the origin at the second zero crossing and consider only rates above that at which the change from loss to profit occurs.

The linear functional that is zero at the vector origin is

$$z = \sum_1^n a_i x_i \qquad [4]$$

This homogeneous function of the first degree satisfies Euler's equation

$$\Sigma x_i \frac{\partial z}{\partial x_i} = z \qquad [5]$$

Using the language of geometry, this equation says that the scalar product of the vector $(x_1 \ldots x_n)$ and the gradient of z is the functional z. We shall make the desired generalization by modifying Equation [5] through the requirement that the scalar product, though remaining proportional to z, decreases in the simplest manner proportionately to the magnitude of the vector. We are thus led to consider functions z that are solutions of the equation

$$\Sigma x_i \frac{\partial z}{\partial x_i} = (1 - \Sigma \alpha_i x_i) z, \quad \alpha_i \geq 0, \quad i = 1, 2, \ldots n \ldots [6]$$

It is shown in the Appendix that the solution of this equation which is a natural generalization of Equation 4 is

$$z = \Sigma a_i x_i e^{-\Sigma \alpha_i x_i} \qquad [7]$$

In order not to clutter the discussions with distracting mathematical deductions the investigation of the properties of Equation [7] will be relegated to the Appendix, where proofs will be found for statements made hereafter.

If one component of the vector, say, x_i, is allowed to increase while the others are kept small, the asymptotic form of Equation [6] in the component x_i may be written in the form

$$dz/z = -\alpha_i dx_i \qquad [8]$$

This is the characteristic equation for high rates of production of any component of the system. In words—the fractional drop in profit, at high rates of production, is proportional to the added rate of production. This statement may perhaps be useful in making an estimate of the α's. The a's of Equation [7] would be determined as in the linear case, Equation [4]; indeed, they are the coefficients in the early linear stage of the phenomenon. The linear factor in Equation [7] dominates at low rates of production; the exponential dominates at high rates.

There is no point in a space of dimensionality greater than unity at which the functional of Equation [7] attains a maximum. It attains maxima only along lines. Along every line in n-space which is not in one of the two characteristic $(n - 1)$-spaces, there occurs a maximum of the function. This means that z can have a maximum only on the lines in which $n - 1$ of the $(n - 1)$-dimensional bounding spaces intersect or at the points of intersection of n of these spaces. This is a very great simplification over the possibilities presented by the general nonlinear functional.

The maximum along a line occurs at the point of intersection of the line and the $(n - 1)$-space

$$\Sigma a_i x_i = \lambda \qquad [9]$$

where λ is a parameter that depends on the direction of the line. If the λ's of the set associated with the lines that are intersections of the constraining spaces are all sufficiently large, there will be no intersections within the region of constraint and the problem is substantially that of maximizing the linear functional Equation [4], even though in the present case the departure from linearity can be great. If some, or all, of the λ's are sufficiently small, then the problem consists of determining the intersections of the lines of the constraining set and the spaces of Equation [9].

Another useful result is that the maxima along radial lines occur in the $(n-1)$-space

$$\Sigma \alpha_i x_i = 1 \quad \ldots\ldots [10]$$

An Example

The features of this model may best be illustrated, and the means of solution indicated, by a simple example; so simple, in fact, that it is unencumbered with the arithmetical apparatus that characterizes the solution of the usual problem.

We consider a 2-component system in which x and y are the rates at which the components, X, Y of a substance are produced. For example, x may be the number of items X produced in one week, or it may be the number of grams of X produced in a day. In the first case x could take only integral values. In the second case the variation would be continuous. As usual we take the variation to be continuous, using the nearest integer value in the discontinuous case.

Let the profit associated with the rate of production of a combination of X and Y be

$$z = (ax + by)e^{-\alpha x - \beta y} \quad \ldots\ldots [11]$$

and let there be four constraints $0 \leq x \leq A$; $0 \leq y \leq A$. In particular, let $a = 4$, $b = 3$, $\alpha = \beta = 0.002$. It should be noted that, if $\alpha = \beta = 0$, the problem is linear

$$z = 4x + 3y, \quad 0 \leq x \leq A, \quad 0 \leq y \leq A \quad \ldots\ldots [12]$$

and quite trivial, for it is apparent that the maximum profit is obtained when both x and y attain their maximum values A, regardless of the magnitude of A. This optimum profit in the linear case is $z = 7A$.

Where saturation is considered the profit is

$$z = (4x + 3y)e^{-0.002(x+y)}, \quad 0 \leq x \leq A, \quad 0 \leq y \leq A \quad [13]$$

To obtain a maximum we differentiate

$$\left.\begin{aligned} \frac{\partial z}{\partial x} &= [4 - 0.002(4x + 3y)]e^{-0.002(x+y)} \\ \frac{\partial z}{\partial y} &= [3 - 0.002(4x + 3y)]e^{-0.002(x+y)} \end{aligned}\right\} \quad \ldots\ldots [14]$$

These derivatives are zero if

$$\left.\begin{aligned} 4x + 3y &= 2000 \\ 4x + 3y &= 1500 \end{aligned}\right\} \quad \ldots\ldots [15]$$

There is no solution common to the two equations since the lines they represent are parallel, from which it follows that there is no maximum. However, $\partial z/\partial x$ is zero along the first line of Equations [15] and $\partial z/\partial y$ is zero along the second. These are special cases of Equation [9]. These lines are shown as the two parallel lines CD and AB in Fig. 4, the arrows indicating the directions in which the maxima occur.

Suppose that $A = 100$. The constraining region is the square in Fig. 4 bounded by the co-ordinate axes and the lines $x = 100$, $y = 100$. The maxima along the lines $y = 0$, $y = 100$ occur where these lines intersect the line CD. These intersections are outside the constraining region, whence the value of z on these lines rises monotonically with x. Also, the maxima on the lines $x = 0$, $x = 100$ occur outside the constraining region since the points are on AB. It follows that the solution is $x = 100$, $y = 100$ as in the linear case. The profit is different in the two cases, however; for substituting these values of x and y in Equation [13] we obtain $z = 469$, whereas substitution in the linear function, Equation [12], yields $z = 700$.

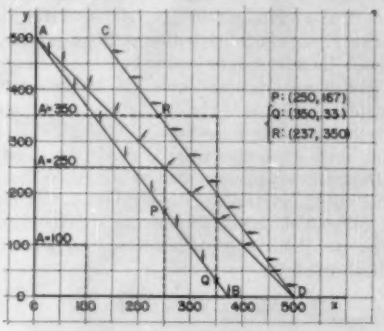


Fig. 4 Occurrence of Maxima on Boundaries

If we now let $A = 250$ we see in Fig. 4 that the line AB intersects the boundary of the constraining region. It is only at the point P that a maximum occurs since it is only the line $x = 250$ that has the proper direction for a maximum at an intersection with this line. Since the value of z at P is a maximum for this line, it is higher than the value at the corner $x = 250$, $y = 250$. But the value of z at this corner is the highest on the line $y = 250$, since the intersection of this line with the line CD occurs outside the proper region. Hence the required solution is $x = 250$, $y = 167$, the co-ordinates of P.

Finally, if $A = 350$, there are two intersections at which maxima occur: $Q = (350, 33)$ on $x = 350$, and $R = (237, 350)$ on $y = 350$. It is now a matter of substituting those pairs of values in Equation [13] to determine the higher of the two.

Since the computations are simple enough in this system the values for all the cases have been determined on the boundaries and are shown in Fig. 5. Equation [13] has been reduced in the following manner: $\xi = x/A$, $\eta = y/A$, $\zeta = z/A$, $\mu = 0.002A$, so that

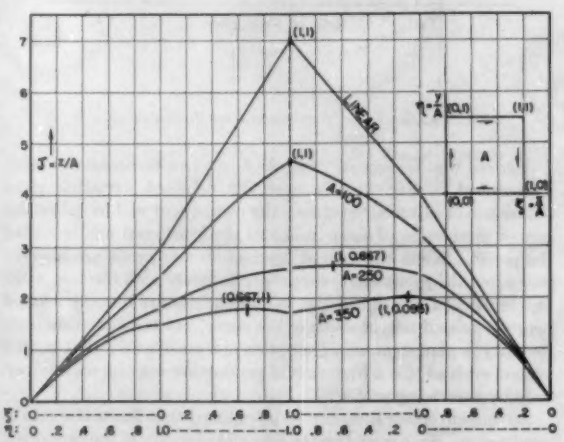


Fig. 5 Values of Functional on Boundary

$$\zeta = (4\xi + 3\eta)e^{-A(\xi+\eta)}, \quad 0 \leq \xi \leq 1, \quad 0 \leq \eta \leq 1 \ldots [16]$$

The curves shown in the figure represent the values of ζ on the bounding square shown in the upper right of the figure, starting at (0, 0) and proceeding as shown by the arrows. The three lower curves are for $A = 100$, 250, and 350, which lead to $\mu = 0.2$, 0.5, 0.7 in Equation [16]. The linear configuration remains fixed for all values of A.

Using the same functional we shall consider the problem without external constraints ($A \to \infty$) in a somewhat different manner. Suppose there are two producers of a product, X producing at the rate x, Y at the rate y. If X were the sole producer his profit would be

$$z_x = 4xe^{-0.002x}$$

and if Y were the sole producer his profit would be

$$z_Y = 3ye^{-0.002y}$$

The coefficient 0.002 in the exponential must be the same for both if the products flow into the same market. If they both produce simultaneously their profits are

$$z_x = 4xe^{-0.002(x+y)}, \quad z_Y = 3ye^{-0.002(x+y)} \ldots\ldots [17]$$

respectively. The exponentials are the same in both cases since $x + y$ is the rate at which the product from both sources flows into the market.

Suppose Y is producing at the rate $y = 400$, when X begins to produce. Let Y continue to produce at the rate $y = 400$. Initially Y's profit is

$$z = 1200e^{-0.8} = 539$$

X may agree to produce at the rate that makes the combined profit a maximum. Maximizing

$$z = (4x + 1200)e^{-0.002(x+400)}$$

and substituting in Equation [17] we obtain for this case

$$y = 400, \qquad z_Y = 361$$
$$x = 200, \qquad z_X = 241$$
$$\text{Combined profit } z = 602$$

X may decide to produce at the rate that will bring him a maximum. In this case the quantity to be maximized is

$$z = 4xe^{-0.002(x+400)} = 1.797xe^{-0.002x}$$

and we have

$$y = 400, \qquad z_Y = 198$$
$$x = 500, \qquad z_X = 331$$
$$\text{Combined profit } z = 529$$

Suppose finally that X and Y combine and agree that the rate of production shall be that which gives the greatest combined profit with the ratio y/x maintained at the constant value m. The function to be maximized is

$$z = (4 + 3m)xe^{-0.002(1+m)x} \ldots\ldots [18]$$

which yields

$$x = \frac{500}{1+m}, \quad y = \frac{500m}{1+m} \ldots\ldots [19]$$

These points fall on the line $x + y = 500$, which is an instance of Equation [10], where it appears in the form $0.002(x + y) = 1$. This is line AD in Fig. 4. Although for a fixed m the maximum is obtained at the point Equation [19], there is no value of m at which an absolute maximum is attained. The line $x + y = 500$ is one of those exceptional lines already mentioned on which no maximum occurs. For three values of m the profit is shown in Table 1.

TABLE 1 PROFIT FOR THREE VALUES OF m

m	x	y	z
0	500	0	736
1	250	250	644
∞	0	500	552

These illustrations indicate that the model presented here may be useful in programming studies in important fields of application where the linear approximation could at best be applied piece by piece. The formal properties of the functional Equation [7] indicate that the maximizing problem may be reasonably amenable to computation. It will be shown in the Appendix that only linear processes are involved in the general problem.

Appendix

A notation will be adopted here that has been used extensively[5] and which avoids unwieldy typography. As used here, a, b, c, α, β, m, x, y are column vectors and the same letters with primes are their transposes, which are row vectors. For example

$$a = \begin{Bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{Bmatrix}, \quad b' = \{b_1 b_2 \ldots b_n\}$$

The scalar product of a and b is

$$a'b = b'a = \Sigma a_i b_i$$

from which it follows that

$$|a|^2 = a'a = \Sigma a_i^2$$

The gradient operator ∇ is the symbolic column vector

$$\nabla = \begin{Bmatrix} \partial/\partial x_1 \\ \vdots \\ \partial/\partial x_n \end{Bmatrix}$$

An n-vector is a vector (or its transpose) with n-components. Thus, if f is a scalar, ∇f is an n-vector.

A is an $n \times r$ matrix with elements A_{ij}. A_c is a column vector.

The symbols f, g, h, u, v will here be used for functionals of x; n, k, and r will be used for integers; and λ, μ for scalar parameters.

The gradient of the homogeneous linear form

$$z = a'x \ldots\ldots [20]$$

is the constant vector $\nabla z = a$. Now, $x'\nabla z = x'a = a'x = z$, which is Euler's equation for a homogeneous form of first degree. If, on the other hand, we start with Euler's equation

$$x'\nabla z = z \ldots\ldots [21]$$

we have, of course, the solution Equation [20], in which a is arbitrary, but this is not the general solution since the general solution of this equation must contain an arbitrary function of the variables $x_1, \ldots x_n$. Let $z = fg$. Then

$$\nabla fg = g\nabla f + f\nabla g$$

and

$$x'\nabla fg = gx'\nabla f + fx'\nabla g$$

This product satisfies Equation [21] if

[5] The notation is used in the paper cited in footnote 2. It is explained in H. Cramér's "Mathematical Methods of Statistics," Princeton University Press, Princeton, N. J., 1951, p. 106, et seq.

$$\left.\begin{aligned} x'\nabla f &= f \\ x'\nabla g &= 0 \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots [22]$$

A particular solution of the first of these, which is the same as Equation [21], is, as we have seen, $f = a'x$. The generality arises from the solution of the second of Equations [22]. Let

$$\theta = x_1^{p_1} x_2^{p_2} \ldots\ldots x_n^{p_n}$$

It is seen that

$$x_i \frac{\partial \theta}{\partial x_i} = p_i \theta$$

Suppose that $g = g(\theta)$. Then

$$x'\nabla g(\theta) = (p_1 + p_2 + \ldots\ldots + p_n)\theta \frac{dg}{d\theta}$$

This is identically zero if

$$p_1 + p_2 + \ldots\ldots p_n = 0$$

Consequently, the general solution of Equation [21] is

$$z = a'xg(\theta) \ldots\ldots\ldots\ldots\ldots\ldots [23]$$

where g is an arbitrary function of degree zero in the variables $x_1 \ldots x_n$. The form Equation [20] that we wish to generalize is obtained from this by letting $g = 1$.

The generalization is made through the equation

$$x'\nabla z = (1 - \alpha'x)z \ldots\ldots\ldots\ldots\ldots\ldots [24]$$

Let $z = uv$. Then

$$x'\nabla uv = vx'\nabla u + ux'\nabla v$$

The equation is satisfied by this product if

$$\left.\begin{aligned} x'\nabla u &= u \\ x'\nabla v &= -\alpha'xv \end{aligned}\right\} \quad \ldots\ldots\ldots\ldots\ldots [25]$$

The first of these has been found to have the general solution $u = a'xg(\theta)$, while it may be verified that the particular solution $v = \exp(-\alpha'x)$ satisfies the second. Whence the general solution of Equation [24] is

$$z = a'xe^{-\alpha'x}g(\theta)$$

and the proper generalization of Equation [20] we obtain by setting $g = 1$. The generalized functional is therefore

$$z = a'xe^{-\alpha'x} \ldots\ldots\ldots\ldots\ldots\ldots [26]$$

Sufficiently near $x = 0$, $\alpha'x$ is negligible compared to unity and Equation [24] reduces approximately to $x'\nabla z \backsimeq z$, which is Equation [21]. For sufficiently large values of x, $1 \ll \alpha'x$ and $x'\nabla z \sim -\alpha'xz$; that is, $\nabla z \sim -\alpha z$.

The gradient of z is explicitly

$$\nabla z = [a - \alpha(a'x)]e^{-\alpha'x}$$

and $|\nabla z|^2 = \nabla' z \nabla z = [a'a - 2a'\alpha(a'x) + \alpha'\alpha(a'x)^2]e^{-2\alpha'x}$

For a maximum $|\nabla z| = 0$. This implies that

$$\alpha'\alpha(a'x)^2 - 2a'\alpha(a'x) + a'a = 0$$

The discriminant of this quadratic is

$$4[(a'\alpha)^2 - (\alpha'\alpha)(a'a)] \leq 0$$

the inequality being that of Schwarz.[5] If $a \neq \mu\alpha$, the discriminant is negative and the functional has no maximum. If $a = \mu\alpha$, then

$$(a'x)^2 - 2\mu(a'x) + \mu^2 = 0$$

and the functional has a stationary value on the whole $(n-1)$-space $a'x = \mu$ equal to μe^{-1}.

The points of the $(n-r)$-space which is the intersection of the r $(n-1)$-spaces

$$Ax = A_0$$

are covered by a proper $(n-r)$-vector y.[4] On this base the functional becomes

$$z = (b_0 + b'y)e^{-(\beta_0 + \beta'y)} \ldots\ldots\ldots\ldots [27]$$

and, so long as $n - r > 1$, this constrained functional has no maximum, as was shown previously. If $n - r = 1$, then z is simply a function of the scalar parameter y and it has a maximum at $y = 1/\beta$ equal to $(b_0 + b/\beta)\exp(-\beta_0 - 1)$.

Let c, such that $c'c = 1$, be an arbitrary direction in n-space. The directional derivative along c is

$$c'\nabla z = [(c'a - c'\alpha(a'x)]e^{-\alpha'x}$$

If $c'a \neq 0$, $c'\alpha \neq 0$, the directional derivative is zero on the $(n-1)$-space

$$a'x = \frac{c'a}{c'\alpha} \ldots\ldots\ldots\ldots\ldots\ldots [28]$$

It follows that the functional has a maximum on a line with direction c at the point where this line pierces the $(n-1)$-space, Equation [28]. If $c'a = 0$, c is in the $(n-1)$-space $a'x = \lambda$, and if $c'\alpha = 0$ it is in the $(n-1)$-space $\alpha'x = \mu$. It is seen that no maxima exist for these exceptional lines.

If $c_j = 0$, $j \neq i$,; $c_i = 1$ the differentiation is along a coordinate direction and the maxima occur on the space $a'x = a_i/\alpha_i$. The i-intercept of this space, obtained by letting $x_j = 0$, $j \neq i$, is

$$x_i = 1/\alpha_i \ldots\ldots\ldots\ldots\ldots\ldots [29]$$

If m is a unit vector $[m'm = 1]$ then $x = m\lambda$ is a radial line with λ the radial distance from the origin. On these lines the functional is

$$z = (a'm)\lambda e^{-(a'm)\lambda}$$

which has a maximum at $\lambda = 1/(\alpha'm)$, and $x = m\lambda = m/(\alpha'm)$. Taking the scalar product of this with α we find

$$\alpha'x = 1 \ldots\ldots\ldots\ldots\ldots\ldots [30]$$

This is the space on which the functional has maxima along the radial lines. This is the space determined by the intercepts Equation [29].

Example

As an illustration we shall take the same functional as that of Equation [12], but we shall use perfectly general processes. Figs. 6 and 7 are 2-space configurations, but one may think of the triple lines as being $(n-1)$-spaces in the generalization, the single lines being ordinary lines (1-spaces) but not necessarily in the plane of the paper.

We wish to find the point $x = M$ at which $z = a'xe^{-\alpha'x}$ is a maximum on the line $\overline{AB} = b + \lambda c$, $c'c = 1$, Fig. 6. Draw $\overline{OC} = \lambda c$. Through the point C at which it intersects the characteristic space $\alpha'x = 1$, pass the space $a'x = \lambda_{AB}$ parallel to the characteristic space $a'x = 0$. This is the space $a'x = a'c/\alpha'c$ of Equation [28]. The point M at which this space is pierced by $\overline{AB}$ is the required point. The line $\overline{OC}$ is called $\overline{ab}$ to indicate its relation to $\overline{AB}$, the given line. The processes used to deter-

[5] Birkhoff and MacLane, loc. cit.,[4] p. 190 theorem 17 (iii).

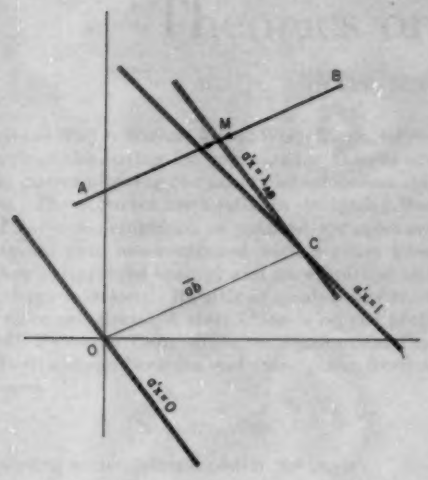


FIG. 6 LOCATION OF MAXIMUM IN A LINE

mine M are all linear. Explicitly

$$M = b + \left(\frac{1}{a'c} - \frac{a'b}{a'c}\right) c$$

From this the maximum value of z is found to be

$$z_{max} = \frac{a'c}{a'c}\, e^{-1-a'b+\frac{a'b}{a'c}\, a'z}$$

In Fig. 7 we consider the boundary O-A-B-C-D-O. The previous construction is repeated for all the 5 lines of the boundary.

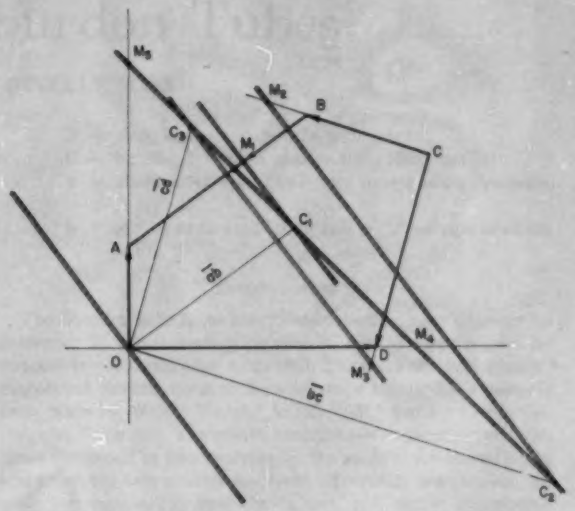


FIG. 7 LOCATION OF OPTIMA ON CONSTRAINING BOUNDARIES

For example, $\overline{cd}$ is parallel to $\overline{CD}$. It intersects the characteristic space $a'x = 1$ at C_3. The triple line C_3M_3 represents the space through C_3 parallel to the characteristic space $a'x = 0$. Since the maximum on the co-ordinate line OA occurs at M_5, the value of the functional continues to increase from A to M_1, where the maximum on AB occurs. Likewise, there is no maximum at the corner B. There is a maximum at the corner D, however, since the maxima on OD and CD are at M_4 and M_3, outside the boundary. C is a minimum point, since the functional rises in both directions.

This example illustrates all the nondegenerate possibilities.

Theories on Bourdon Tubes

By F. B. JENNINGS,[1] BEVERLY, MASS.

Theories of Walter Wuest; Alfred Wolf; Clark, Gilroy, and Reissner; and the author are compared. Results are presented in curves plotting the same dimensionless ratios in all cases. These curves are useful in designing Bourdon tubes of flat-oval, elliptical, or pointed-arc cross section. Experimental data are compared with a curve based on the author's simplified theory, and an empirical curve of similar shape is drawn. Results of analyses of tip travel and tip force are given. A start is made on the project of the ASME Research Committee on Mechanical Pressure Elements to change Bourdon-tube designing from an art to a science.

Nomenclature

The following nomenclature is used in the paper:

- A = constant in bulging function, Equation [5]
- a = half-width of tube cross section, Fig. 2.
- α_1 and α_2 = principal directions for tip force, Equations [28] and [30]
- B = constant in bulging function, Equation [6]
- b = half-height of tube cross section, Fig. 2
- β = parameter in bulging function, Equation [7]
- c = quarter-circumference of flat-oval section, Equation [23]
- $\gamma = 2\beta a$ = parameter in rectangular tube theory
- Δ = increase in operand, $-\Delta$ = decrease
- E = Young's modulus
- F, F_1, F_2 = tip force and its components in principal directions, Fig. 14
- $f_1 = \Delta R_p E b h^3 / R p a^4$ = pressure sensitivity ratio, Equation [9]
- $f_2 = MR^2/EI\Delta R_M$ = stiffness ratio, Equation [10]
- $f_3 = S/(pa^2/h^2 + Mba^2/IRh)$ = stress ratio, Equation [12]
- ϕ = direction of tip travel, Fig. 11
- $G = |G|e^{i\not\preceq G}$ = complex function of θ, Equation [31]
- g_1, g_2 = principal flexibility ratios, Equations [27] and [29]
- $H = H(\vartheta, \gamma)$ = function of ϑ and γ in rectangular tube theory, Equation [20]
- h = wall thickness, Fig. 2
- θ = angular length, Fig. 11
- I = geometrical moment of inertia of cross section
- $\vartheta = b/a$ = height-to-width ratio of tube section
- l = distance of tip travel, Fig. 11
- l_1, l_2 = tip-travel components in principal directions, Fig. 14
- M = bending moment applied to tip of tube in straightening direction
- $\mu = 1/3$ = Poisson ratio
- p = differential pressure, internal minus external
- R = radius of curvature of tube center line, Fig. 2
- S = maximum stress in tube wall
- V = bulging deflection of tube wall, Equation [4]
- x = abscissa of tube wall = independent variable, Fig. 2
- $y = y(x)$ = ordinate of tube wall = tube-shape function, Fig. 2

Introduction

The Bourdon tube is the most widely used primary detector for pressure. It is rather shocking that it is still being designed by cut-and-try methods after more than 100 years of development.[2] Significant contributions to Bourdon-tube theory have recently been made by Walter Wuest,[3] Alfred Wolf,[4] and by Clark, Gilroy, and Reissner.[5] The results brought out in these papers, also those developed in two analyses by the author, are co-ordinated and presented here in graphical form. To clarify comparison, the same nomenclature is used throughout, and curves are plotted against the same dimensionless ratios in all cases. This makes these theories simple to apply in practical design work.

Results, rather than detailed derivations, are stressed throughout the paper. Curves are used extensively because with them the physical situation and the meaning of the equations are more easily visualized. An attempt is made to keep complicated mathematics to a minimum.

The Bourdon-tube theories discussed here give the change in radius of curvature due to application of internal pressure. They also give the change in radius resulting from a pure bending moment acting on the tip of the tube. These are combined when the tube operates against a restraining spring. The maximum stress which results from pressure and bending is also determined because it is of some importance in determining the life of the tube.

In practical applications of Bourdon tubes, the engineer is interested in the distance the tip travels, the angle through which it turns, and the effect of external forces acting on the tip. These practical aspects have been analyzed by the author. Results are presented in equations and curves with a general discussion of pertinent factors.

Wuest's Theory

In this section we present a physical picture of how Bourdon tubes operate based on the theory of W. Wuest. His analysis is approximate, and in particular it is restricted to tubes whose cross sections have small height in comparison with width. It is also assumed that in a cross-section view the wall is everywhere nearly parallel to its width axis. Wuest's theory is well suited to such shapes as the pointed-arc and the "idealized" section shown in Fig. 1. The flat-oval, elliptical, or rectangular shapes cannot be handled accurately by this theory because the walls at the ends are at 90 deg to the walls at the long sides.

The operation of a Bourdon tube may be explained briefly as follows: When pressure is applied inside the tube, the walls bulge

---

[1] President of Ipswich Boat Company; formerly with General Electric Company. Mem. ASME.

Contributed by the Research Committee on Mechanical Pressure Elements and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 7, 1954. Paper No. 54—A-168.

[2] "Bibliography on Bourdon Tubes and Bourdon Tube Gages," by L. M. Van der Pyl, ASME Paper 53—IRD-1, presented at the Eighth National Instrument Conference, Chicago, Ill., September, 1953.

[3] "Der Einfluss der Querschnittsform auf das Verhalten von Bourdonfedern," by W. Wuest, *Ingenieur-Archiv*, vol. 20, 1952, pp. 116–125, art. 1798, translated by G. E. Main Library, Schenectady, N. Y.

[4] "An Elementary Theory of the Bourdon Gage," by Alfred Wolf, *Journal of Applied Mechanics*, Trans. ASME, vol. 68, 1946, p. A-207.

[5] "Stresses and Deformations of Toroidal Shells of Elliptical Cross Section," by R. A. Clark, T. I. Gilroy, and E. Reissner, *Journal of Applied Mechanics*, Trans. ASME, vol. 74, 1952, pp. 37–48.

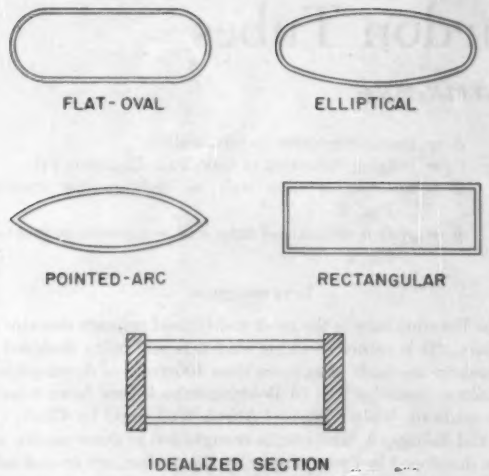


FIG. 1 SHAPES OF BOURDON-TUBE CROSS SECTIONS CONSIDERED IN THIS PAPER

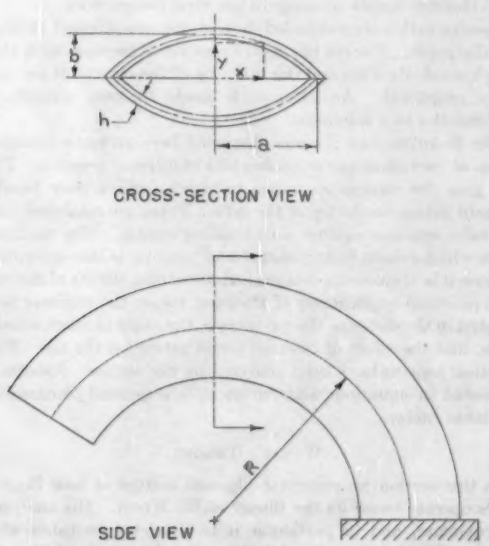


FIG. 2 DRAWING OF BOURDON TUBE SHOWING NOTATION

The outer wall is stretched like a string by this bulging and pulls on the outer edges of the tube ends. The inner wall is compressed like an arch by the bulging, and it pushes on the inner edges of the tube ends. This applies bending moments to the ends of the tube which tend to straighten the coil. The tube deflects in the same way all along its length provided one ignores end effects caused by the fittings which close the tube ends.

The key to Wuest's Bourdon-tube theory is the bulging deflection V. The differential equation relating V to the deflection $\Delta R/R$ and the pressure p is

$$\frac{Eh^3/12}{1-\mu^2}\frac{d^4V}{dx^4} = -\frac{Eh}{R^2}\left(V - y\frac{\Delta R}{R}\right) + p \quad \ldots\ldots [1]$$

The symbols in this equation are defined in Fig. 2 and the nomenclature. The term on the left equals the second derivative with respect to x of the moments which cause the wall to bulge. It also equals the first derivative of forces tending to shear the wall. The first term on the right equals $-h/R$ times the stress along the tube. This term gives the radial-force component caused by the stretching of the ribbonlike element dx of the tube wall. The last term is differential pressure, internal minus external. If one were to derive an accurate differential equation applying to any cross-section shape, it would be much more complicated than Equation [1]. The other three derivatives of V would appear, and there would be a number of variable coefficients. A complete differential equation of this sort becomes rather impractical to solve.

Equation [1] is a very useful first approximation. The solution is subject to boundary conditions imposed by symmetry of the tube section; that is, V is an even function of x, is zero at $x = a$, and has zero slope at $x = 0$ and a.

The solution of Equation [1] provides a bulging function V which depends both on the amount of unbending $\Delta R/R$ and the pressure p. The moment applied to the end of the tube is calculated by suitably integrating the lengthwise stress

$$\left.\begin{aligned} M &= -4\frac{Eh}{R}\int_0^a \left(V - y\frac{\Delta R}{R}\right) y dx \\ &= -4\frac{Eh}{R}\int_0^a V y dx + \frac{EI\Delta R}{R^2} \end{aligned}\right\} \quad \ldots\ldots [2]$$

The conventional formula for effect of bending moment is

$$M = EI\Delta R/R^2$$

and the other term shows that bulging reduces the moment required to get a desired deflection. As V depends on both $\Delta R/R$ and p, this equation with $M = 0$ can also give the sensitivity of a Bourdon tube when no moment is applied to the tip.

The maximum stress ordinarily occurs at the inner surface of the tube wall at the end of the tube section. This stress is proportional to the second derivative of bulging

$$S = \frac{Eh/2}{1-\mu^2}V''(a) \quad \ldots\ldots [3]$$

IDEALIZED SECTION

These equations apply to any shape of tube section provided the height-to-width ratio b/a is small. The simplest case to which Wuest has applied them is the idealized section shown in Fig. 1. In this case y is constant, which considerably simplifies the solution. The cross-hatched ends are assigned special properties which need explanation. They are infinitely rigid to all stresses except unbending—to which they are perfectly flexible. This means that all bulging occurs in the horizontal walls and that the boundary conditions previously described hold in this case. This analysis applies particularly to flat-oval sections which are very much wider than their height.

Subject to these assumptions and boundary conditions, the solution of Equation [1] for the idealized section is

$$V = \left(b\frac{\Delta R}{R} + \frac{R^2p}{hE}\right)(1 + A\cosh\beta x\cos\beta x + B\sinh\beta x\sin\beta x) \quad \ldots\ldots [4]$$

where

$$A = -\frac{\sinh\beta a\cos\beta a + \cosh\beta a\sin\beta a}{\cosh\beta a\sinh\beta a + \cos\beta a\sin\beta a} \quad \ldots\ldots [5]$$

$$B = \frac{\sinh\beta a\cos\beta a - \cosh\beta a\sin\beta a}{\cosh\beta a\sinh\beta a + \cos\beta a\sin\beta a} \quad \ldots\ldots [6]$$

$$\beta = \frac{3^{1/4}(1-\mu^2)^{1/4}}{\sqrt{Rh}} \quad [7]$$

The presence of the cosine and sine terms in this bulging function indicates that the wall deflection has an oscillatory character if the tube width is large enough. A physical explanation for this waviness is that longitudinal tension in the walls vies with lateral bending moment. One important result is that, when the radius of curvature of the Bourdon tube is sufficiently large, the oscillatory effects do not show up. This is a physical explanation for the two slopes seen in the characteristic curves of Figs. 3 to 10.

The deflection of the idealized Bourdon tube, when both pressure and bending-moment loads are applied, is obtained by carrying out the integration called for in Equation [2]

$$\frac{\Delta R}{R} = \frac{f_1 a^4}{bh^3 E} p + \frac{R}{f_2 EI} M \quad [8]$$

$$f_1 = \frac{R^2h^2}{a^4}\left(\frac{1}{f_2} - 1\right) \quad [9]$$

$$f_2 = \frac{1}{\beta a}\frac{\cosh 2\beta a - \cos 2\beta a}{\sinh 2\beta a + \sin 2\beta a} \quad [10]$$

The maximum stress in the idealized tube is obtained by taking the second derivative of the bulging function

$$S = f_3\left(\frac{a^2}{h^2} p + \frac{a^2 b}{Rh\, I} M\right) \quad [11]$$

$$f_3 = \frac{3}{\beta a}\frac{\sinh 2\beta a - \sin 2\beta a}{\cosh 2\beta a - \cos 2\beta a} \quad [12]$$

From this equation one can calculate the stress when pressure, bending moment, or both are applied. It is important to realize that the stress is lower if the tip is restrained than if it is free. When the tip is fixed, $\Delta R/R = 0$, and the restraining moment M calculated from Equation [8] is negative. In Equation [11] the moment term subtracts from the pressure term resulting in a smaller stress value. This is the basis for the common practice of having a Bourdon tube hit stops which provide some protection when excessive pressures are applied.

It is convenient to define dimensionless ratios based on Equations [8] and [11]. The ratios we have chosen to use in plotting curves are as follows: First, a tube parameter related to βa is defined:

Tube Parameter

$$\frac{Rh}{a^2} = \frac{\sqrt{3(1-\mu^2)}}{(\beta a)^2} \quad [13]$$

This tube parameter can be considered a measure of the radius of curvature of the Bourdon tube. It is convenient to set Poisson's ratio $\mu = 1/3$ in all numerical work. This is a reasonable approximation because $1 - \mu^2$ does not change much over the range of μ-values encountered with different materials.

The pressure-sensitivity ratio is defined by setting bending moment $M = 0$ in Equation [8] and solving for the function f_1:

Sensitivity Ratio

$$f_1 = \frac{\Delta R_p}{R}\frac{E}{p}\frac{bh^3}{a^4} \quad [14]$$

The stiffness ratio is obtained from Equation [8] by setting $p = 0$ and solving for the function f_2:

Stiffness Ratio

$$f_2 = \frac{MR^2}{EI\Delta R_M} \quad [15]$$

The stress ratio is the function f_3 in Equation [11]. Usually one calculates the stress in an unrestrained Bourdon tube in which the applied moment M is zero. In this case we have

Stress Ratio

$$f_3 = \frac{S_p}{p}\frac{h^2}{a^2} \quad [16]$$

These functions are given for the idealized section by Equations [9], [10], and [12] and are plotted in Figs. 3 and 4. The

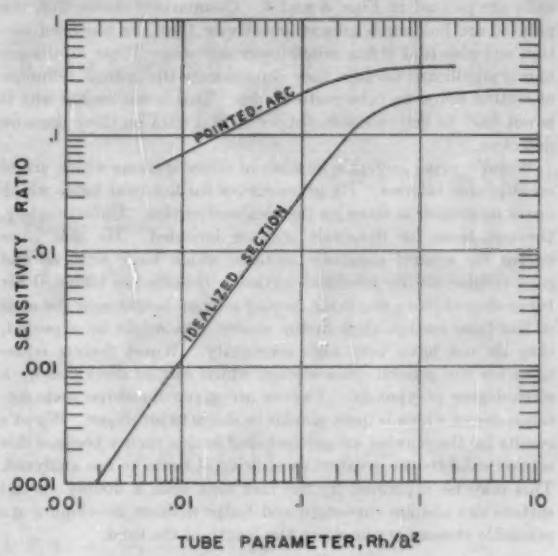


FIG. 3 WUEST'S SENSITIVITY-RATIO CURVES FOR IDEALIZED SECTION AND POINTED-ARC SECTION
(These curves apply for all height-to-width ratios.)

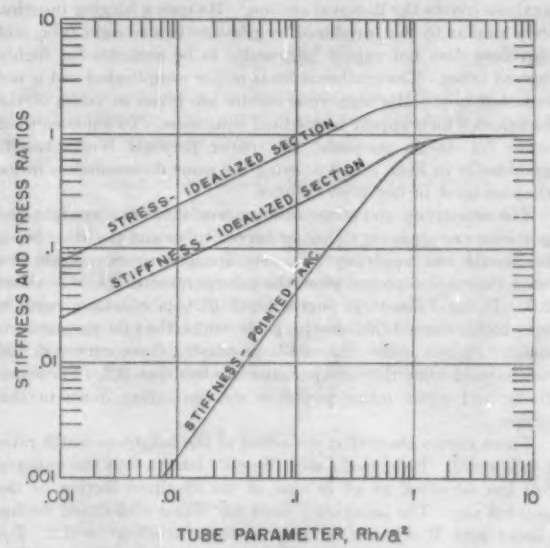


FIG. 4 WUEST'S STIFFNESS-RATIO CURVES FOR IDEALIZED SECTION AND POINTED ARC, AND THE STRESS-RATIO CURVE FOR IDEALIZED SECTION
(These curves apply for all height-to-width ratios.)

curves show that the various functions level off as the radius of curvature becomes large and the tube approaches straightness. The transition region where the curves change slope occurs when the angle $2\beta a$ in the sine and cosine functions is around 180 deg. One could say that the transition occurs when the width of the tube is about a wave length. Many successful commercial Bourdon tubes fall in this transition region.

Wuest has analyzed a number of different sections by the methods which have been outlined in the foregoing. Among these is the pointed arc made up of two parabolas, for which results are plotted in Figs. 3 and 4. Comparison shows that the pointed arc has much greater sensitivity than the idealized section and also that it has much lower stiffness. These results are highly significant because they demonstrate the critical influence of section shape on tube performance. This is one reason why it is not easy to get consistent experimental data on these pressure detectors.

Wuest's paper covers a number of other sections which are of considerable interest. He gives curves for flat-oval tubes which are quite similar to those for the idealized section. Unfortunately, the equations for flat-ovals are not included. He also gives curves for several parabolic sections which have artificial end caps similar to the idealized section. Results are included for tubes shaped like a dog bone, having a larger height near the ends of the tube section than in the center. As might be expected, they do not have very high sensitivity. Wuest derives equations for the general cross section which can be described by a sixth-degree polynomial. Curves are given for three such sections, one of which is quite similar in shape to an ellipse. Wuest's results for the pointed arc are included in this review because this section exhibits the greatest sensitivity of those he has analyzed. This may be explained by the fact that such a doubly curved surface can change curvature and bulge without developing appreciable stresses acting down the length of the tube.

Wolf's Theory

Alfred Wolf applies energy methods and his theory is based on assuming an approximately correct bulging function. His analysis covers the flat-oval section. He uses a bulging function very similar to that produced by pressure in a straight tube, and therefore does not expect his results to be accurate for highly curved tubes. The mathematics is rather complicated and is not repeated here. His numerical results are given in tables of the functions which appear in his final equations. To enhance their value for design purposes, this paper presents Wolf's results graphically in Figs. 5 and 6, using the same dimensionless ratios that are used for the other theories.

The sensitivity and stress ratio curves show two straight-line portions, one applying to highly curved tubes and the other being horizontal and applying to nearly straight tubes. Again the transition region occurs when the tube parameter Rh/a^2 is about 0.5. The stiffness-ratio curves level off to a constant value in very highly curved Bourdon tubes, for which the tube parameter is small. As this is not reasonable physically, these curves should not be used when the tube parameter is less than 0.2. The sensitivity and stress ratios probably also go astray down in this region.

These curves show that the effect of the height-to-width ratio b/a is small. This checks with Wuest's results that the value of b/a has no effect at all in case of the idealized section or the pointed arc. The sensitivity ratio for Wuest's idealized section checks with Wolf's $b/a = 0$ curve down to $Rh/a^2 = 0.2$. For values of the tube parameter below 0.2 the curves diverge. The corresponding curves for stiffness and stress do not check as well, but they do level off to the same values for tubes with large radii.

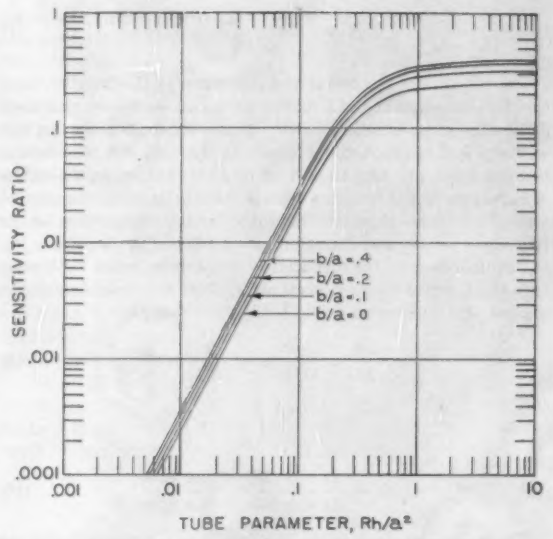


Fig. 5 Sensitivity-Ratio Curves for Flat-Oval Sections Plotted From Wolf's Theory

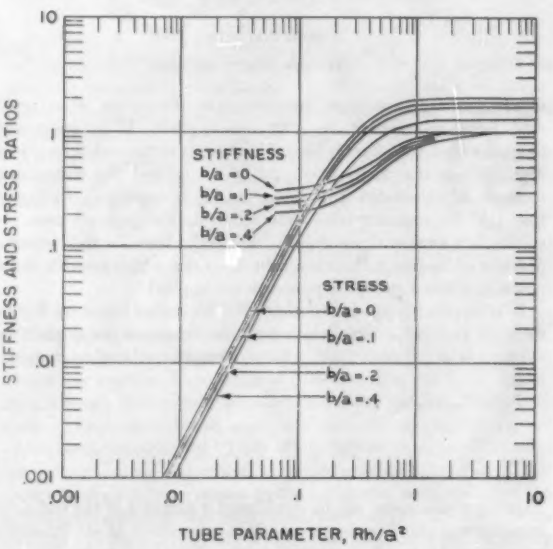


Fig. 6 Stiffness and Stress-Ratio Curves for Flat-Oval Sections Plotted From Wolf's Theory

Clark, Gilroy, and Reissner's Theory

The analysis by Clark, Gilroy, and Reissner of Bourdon tubes with elliptical cross sections utilizes Fourier series and asymptotic approximation methods. The mathematics is complicated and is not reproduced here. Results are plotted in Figs. 7 and 8 using the same dimensionless ratios so that this theory can easily be compared with the others. Some inaccuracies may have been introduced by scaling the rather small curves reproduced in the original paper. For example, we could not determine accurately where the sensitivity-ratio curves leveled off for tubes of low curvature. The sensitivity of the elliptical tubes lies between

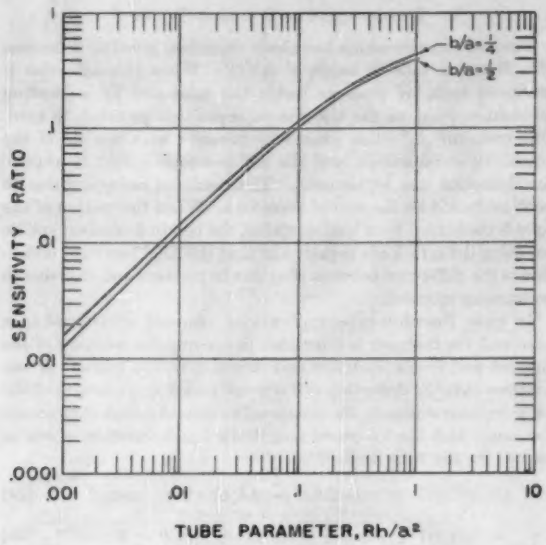


FIG. 7 SENSITIVITY-RATIO CURVES FOR ELLIPTICAL SECTIONS PLOTTED FROM CLARK, GILROY, AND REISSNER'S THEORY

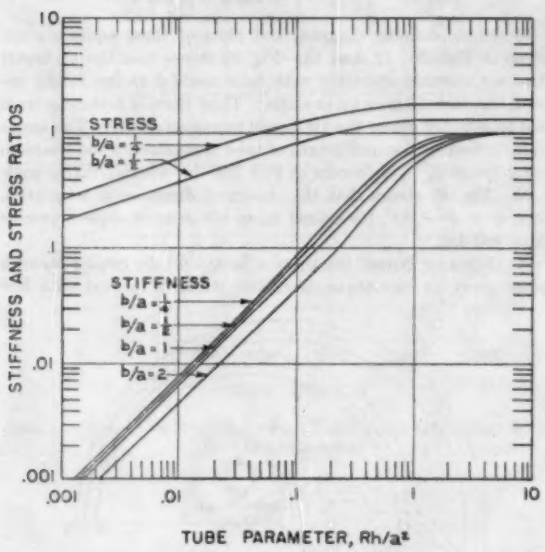


FIG. 8 STIFFNESS AND STRESS-RATIO CURVES FOR ELLIPTICAL SECTIONS PLOTTED FROM CLARK, GILROY, AND REISSNER'S THEORY

those of pointed-arc tubes and idealized or flat-oval tubes. The curves for elliptical tube stress and stiffness ratio check approximately with Wuest's and Wolf's curves for large values of the tube parameter. For highly curved Bourdon tubes, this theory indicates the stress to be higher and the stiffness lower for elliptical tubes than for the two straight-sided tubes. However, study of both the sensitivity and stress curves shows that the elliptical section should give a greater deflection per unit stress than the straight-sided tubes. Ordinarily, this is desirable.

RECTANGULAR TUBE THEORY

The author has made an analysis for tubes of rectangular cross section. The bulging of the horizontal walls is similar to Equation [4] which was derived for the idealized section. The deflection of the vertical walls must satisfy a different differential equation. The solutions are fitted together at the corners of the rectangle by matching slopes and curvatures. This shape is of limited practical value because of undesirable stress concentrations at the four sharp corners. However, the analysis gives an accurate solution. In this case there is no restriction on the value of b/a; but b/R should be small just as is required in the other theories. The equations are as follows:

Sensitivity Ratio

$$\frac{\Delta R_p}{R}\frac{E}{p}\frac{bh^3}{a^4} = \frac{R^2h^3/a^4}{H(\vartheta,\gamma)}\left[\left(\gamma - \vartheta\gamma - \frac{1}{6}\vartheta^2\gamma^2\right)\sinh\gamma + \left(\gamma - \vartheta\gamma + \frac{1}{6}\vartheta^2\gamma^2\right)\sin\gamma + (-2 + \vartheta\gamma^2)\cosh\gamma + (2 + \vartheta\gamma^2)\cos\gamma\right] \quad [17]$$

Stiffness Ratio

$$\frac{MR^2}{EI\Delta R_M} = H(\vartheta,\gamma)/\gamma\left(1 - \frac{1}{3}\vartheta\right)\{\sinh\gamma + \sin\gamma + \vartheta\gamma(\cosh\gamma + \cos\gamma)\} \quad [18]$$

Stress Ratio

$$S/(pa^2/h^2 + Mba^2/IRh) = [12/\gamma H(\vartheta,\gamma)]\left(1 + \frac{1}{3}\vartheta\right)(\sinh\gamma - \sin\gamma) \quad [19]$$

where

$$H(\vartheta,\gamma) = \frac{4}{3}\vartheta\gamma\,(\sinh\gamma + \sin\gamma) + \left(2 + \frac{1}{3}\vartheta^2\gamma^2\right)\cosh\gamma + \left(-2 + \frac{1}{3}\vartheta^2\gamma^2\right)\cos\gamma \quad [20]$$

$$\vartheta = b/a \quad [21]$$

$$\gamma = 2\beta a \quad [22]$$

The sensitivity ratio for height-to-width ratio $\vartheta = 0.3$ is plotted in Fig. 9. This curve checks remarkably closely with Wolf's results for the flat-oval section.

SIMPLIFIED APPROACH AND EMPIRICAL CURVE

Analyses of the idealized section were made independently by Wuest and the author. The author's approach, aimed at applying the idealized section theory to flat-oval tubes, effectively eliminates height-to-width ratio as an independent variable. In this simplified theory the idealized section is wider than the flat-oval section which it represents. To make the active length of wall the same in both sections, the half-width of the idealized section is taken as a quarter of the "circumference" of the flat-oval

$$c = a + (\pi/2 - 1)b \quad [23]$$

where

a = half-width of flat-oval section
b = half-height of flat-oval section
c = quarter-circumference of flat-oval section
= half-width of corresponding idealized section

This simplified theory is the same as that of the idealized sec-

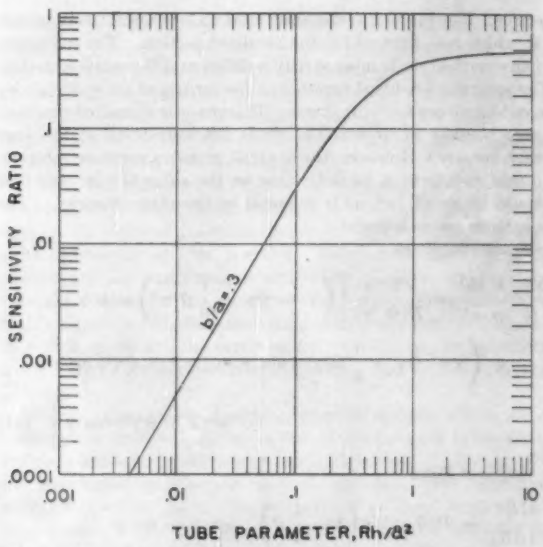


FIG. 9 SENSITIVITY-RATIO CURVE FOR RECTANGULAR SECTION PLOTTED FROM AUTHOR'S THEORY

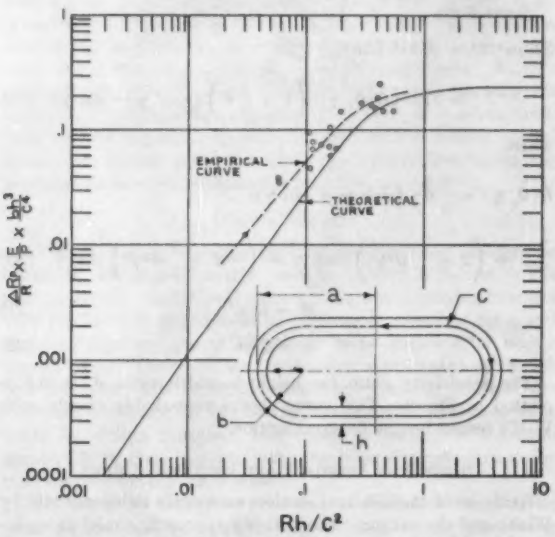


FIG. 10 COMPARISON OF PRESSURE-SENSITIVITY CURVE FOR FLAT-OVAL TUBES, PLOTTED FROM AUTHOR'S SIMPLIFIED THEORY, WITH EXPERIMENTAL DATA AND AN EMPIRICAL CURVE

tion except that a is replaced by c in Equations [8 to 16] and Figs. 3 and 4. There is a single curve independent of b/c for each of the three tube characteristics—sensitivity, stiffness, and stress. Each of these represents a family of curves dependent on b/a in similar manner to the curves given by Wolf's theory. Fig. 10 shows the single theoretical curve for sensitivity ratio, compared with some experimental results. An empirical curve similar in shape to the theoretical curve is drawn through the scatter band of experimental points. The fact that some experimental points are off the empirical curve by a factor as great as 1.5 indicates the need for more accurate test data. Empirical curves of this sort have been found quite useful and accurate for predicting results with tubes of similar shape and manufacture.

DEFLECTIONS OF BOURDON TUBES

The five theories which have been described give the deflection of a Bourdon tube in terms of $\Delta R/R$. When the deflection is produced both by pressure inside the tube and by a bending moment applied to the tip, the superposition principle is used. One finds the deflection caused by pressure with the aid of the sensitivity-ratio curves, and the stiffness-ratio curves determine the deflection due to moment. The resultant deflection due to both loads will be the sum of these two. When the motion of the tube is restrained by a torsion spring, the bending moment and its resulting deflection are negative so that the final resultant deflection is the difference between that due to pressure and that due to restraining moment.

In most Bourdon-tube applications, one end is fastened to a base and the designer is interested in the angular rotation of the tip end and in the distance and direction of tip travel. If one assumes that the deflection ΔR is small and that the length of the tube remains constant, the fractional angle and radius changes are the same, and the tip travel magnitude l and direction ϕ can be related to the tube angle θ

$$-\Delta\theta/\theta = \Delta R/R \quad [24]$$

$$\frac{l}{\Delta R} = \sqrt{(-1 + \cos\theta + \theta\sin\theta)^2 + (\sin\theta - \theta\cos\theta)^2} \quad [25]$$

$$\phi = \tan^{-1}\frac{\sin\theta - \theta\cos\theta}{-1 + \cos\theta + \theta\sin\theta} \quad [26]$$

A symbol-defining diagram and plots of these equations are given in Figs. 11, 12, and 13. Fig. 12 shows that the tip travel does not increase smoothly with tube angle θ as one might expect, but instead goes up in steps. Thus there is a plateau from 300 to 400 deg where the tip travel increases slowly. The maximum deflection per unit length of tube is obtained when the ratio l/θ is greatest. This occurs at $\theta = 240$ deg where $l/\theta\Delta R$ equals 1.26. Fig. 13 shows that the tip-travel direction is tangential, with $\phi = \theta - 90°$, for helical tubes which have angle θ greater than 400 deg.

As shown by Wuest[8] the tip of a Bourdon tube rotates about a polar point, or instantaneous center, which is located on a line

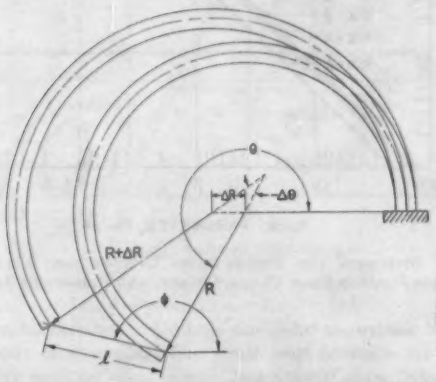


FIG. 11 ILLUSTRATION OF BOURDON-TUBE TIP TRAVEL RESULTING FROM CHANGE IN RADIUS OF CURVATURE

drawn perpendicular to the tip travel l at the end of the radius vector R. The distance from the tip of the tube to the polar point is $Rl/\theta\Delta R$. A gear segment, attached to the tip of a Bourdon tube, with its shaft and bearings located at the polar point puts no

[8] "Die Bewegungslehre von Röhrenfedern," by Walter Wuest, *Zeitschrift für Instrumenten Runde*, vol. 63, 1943, pp. 416–428.

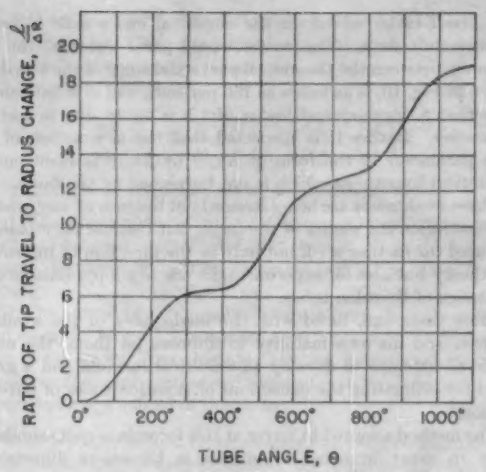


FIG. 12 EFFECT OF TUBE ANGLE ON RATIO OF TIP-TRAVEL DISTANCE TO RADIUS CHANGE

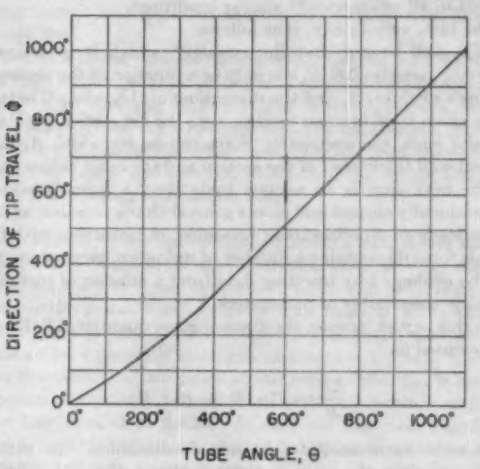


FIG. 13 EFFECT OF TUBE ANGLE ON TIP-TRAVEL DIRECTION WHEN RADIUS IS CHANGED

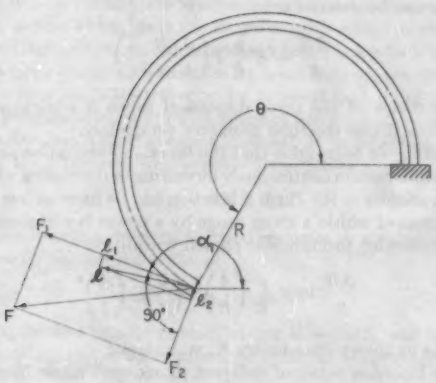


FIG. 14 ILLUSTRATION OF DEFLECTION l PRODUCED BY FORCE F ACTING ON TIP OF BOURDON TUBE

restraint on the tube because the shaft coincides with the axis about which the tube naturally tends to rotate.

EFFECT OF TIP FORCE

It is quite common to apply force to the tip of a Bourdon tube or to restrain the tip to move along a particular line. Equations and curves relating force and tip travel are given later and in Figs. 14 and 15. The situation here is analogous to that encountered in the elastic theory of principal stresses and strains. There are two "principal" directions, 90 deg apart, for which the force and resulting deflection are collinear. If the force is applied in any other direction, the deflection skews off at an angle. In the general case one resolves the force vector into components in the two principal directions. Corresponding deflections are found by multiplying each component by the appropriate "principal" flexibility, and adding vectorially to get the resultant deflection vector. The larger principal flexibility ratio g_1 and its direction α_1 are given as follows

$$g_1 = \frac{l_1}{F_1}\frac{EIf_2}{R^3} = \theta - \sin\theta + |G| \quad [27]$$

$$\alpha_1 = \frac{1}{2} \measuredangle G \quad [28]$$

The smaller principal flexibility ratio and its direction are

$$g_2 = \frac{l_2}{F_2}\frac{EIf_2}{R^3} = \theta - \sin\theta - |G| \quad [29]$$

$$\alpha_2 = \frac{1}{2} \measuredangle G + 90 \quad [30]$$

where

$$G = |G|_e \measuredangle G = -\frac{1}{2}\theta e^{j2\theta} - \frac{3}{4} j e^{j2\theta} + j e^{j\theta} - \frac{1}{4} j \quad [31]$$

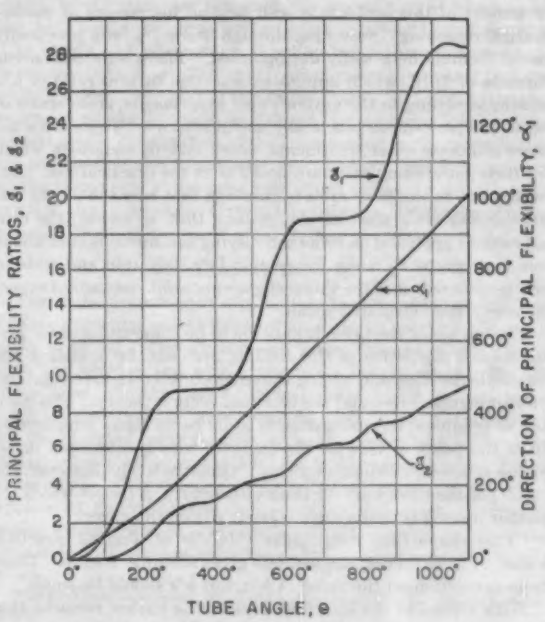


FIG. 15 PRINCIPAL FLEXIBILITY RATIOS AND THEIR DIRECTION PLOTTED AGAINST TUBE ANGLE

The curves plotted in Fig. 15 from these equations show (*a*) the larger principal flexibility is about three times the smaller; (*b*) the flexibilities go up in steps with increasing tube angle; and (*c*) the principal direction α_1 is almost exactly the same as tip-travel direction ϕ given in Fig. 13 for applied moment or pressure. These flexibility ratios are useful either for calculating the deflection when the force is known or for calculating the force when the deflection is given. They may be used to predict performance of a Bourdon tube when its tip is constrained by a linkage to travel in a certain direction. The effects of pressure, applied moment, and applied force on the tip of a Bourdon tube can be superposed to give the resultant deflection in any general case.

Conclusions

The Bourdon-tube designer can obtain much useful information about flat-oval, elliptical, and pointed-arc cross sections from the curves presented in this paper. With their assistance he can change dimensions with good assurance of obtaining desired results, instead of the opposite as sometimes has happened. The author's simplified theory for the flat oval is compared with experimental data. Although data and theory do not check precisely, a useful empirical curve of the same general shape as the theoretical curve is drawn. Results of analyses on tip travel and tip force along with design curves are presented. With all this information at his disposal, the Bourdon-tube designer is in a much stronger position to evaluate experimental data and to develop new tubes. A start is now made on the project of the ASME Research Committee on Mechanical Pressure Elements to change Bourdon-tube designing from an art to a science.

Discussion

K. Gottein.[7] The engineer who is confronted with the task of estimating the deflection of a Bourdon tube or designing a Bourdon tube which would produce a desired deflection when subjected to a given pressure soon finds out that the analytical treatment of this problem is well beyond his powers of mathematical reasoning. Searching through literature for a practically useful formula he is sadly disappointed. There is the old Lorenz formula of 1910, which applies to a section of zero ratio of b/a which, according to the author's own experiments, gives errors of up to 150 per cent on practically useful sections. There are some more elaborate exact treatments which end in equations which by their sheer complexity are prohibitive for practical use. But even if one takes the trouble of thinking that here is now the real answer, one finds after careful reading that, of course, the considerations are based on some simplifying assumptions, and unless one is prepared to make the section infinitely thin and wide or sharp-cornered, or of the thinnest gage material, one cannot expect to verify the calculated result.

The author of the present paper is to be congratulated for giving a lucid exposition of this baffling problem, for collating the available theories, and adding the results of his own investigations on rectangular tubes and the idealized-section theory. The manner of graphical representation is to be particularly commended from the point of view of the engineer who is interested in arriving at a rough estimate quickly. Quantitatively, however, the results of the theory are far from satisfactory, a fact of which the author himself is well aware. Quoting from the paper:

"This shape (the rectangular tube) is of limited practical value. . . . However, the analysis gives accurate results. There is no restriction on the value of b/a, but b/r should be small."

With regard to the idealized section, the author remarks that the spread factor of 1.5 on the empirical curve calls for more accurate test data. The writer would point out that the discrepancy between the theoretical curve and some of the test data, shown in Fig. 10, is as much as 150 per cent, and it is hard to believe that a curve-spread factor of 1.5 is solely due to test inaccuracies. Rather it is suspected that the assumption of the tube parameter in the form of Rh/c^2 or Rh/a^2 is itself only a theoretical hypothesis which is not borne out by the facts.

These weaknesses are being stressed not because of any wish to underestimate the merits of the paper but because they point to the need for further work not only in the direction of improving the theory but also for experimental work to supply reliable data for design of Bourdon tubes.

Some years ago, faced with the inadequacy of the available theories and his own inability to improve on them, the writer made an attempt to develop an empirical formula and a graph sheet for estimating the deflections of Bourdon tubes of flat-oval section.[8]

The method adopted to arrive at this formula is quite similar to what in many branches of physics is known as dimensional analysis. It is basically a sort of curve-fitting process with the advantage which it shares with any theoretical formula that, once proved for a set of conditions, automatically it can be extended to all geometrically similar conditions.

The idea, very briefly, is as follows:

It is well known that the quantity, which it is desired to evaluate, namely, $\Delta R/R$, can only be a function of the pressure p, Young's modulus E, and the dimensions of the tube. The latter, for a given type of cross section, e.g., the flat section with semicircular ends, are completely characterized by width A, height B, and wall thickness t of the section and the bend radius R.

The next step is to assume tentatively a formula which is dimensionally correct and shows general characteristics, as far as is known from experiments or reasoning, in agreement with facts. If this formula contains a number of unknown parameters, these can be evaluated by inserting data from a number of test results equal to the number of parameters.

In this particular case, the dimensionless quantity $\Delta R/R$ can be represented as

$$\frac{\Delta R}{R} = f(p, E, A, B, t, R)$$

In order to make the right side dimensionless, the variables must appear in the form of dimensionless ratios. It is furthermore known that $\Delta R/R$ is proportional to the pressure and inversely proportional to Young's modulus. Consequently, the formula can be written as

$$\frac{\Delta R}{R} = K \frac{p}{E} f\left(\frac{A}{B}, \frac{R}{t}, \frac{A}{t}\right)$$

(The choice of the three dimension ratios is irrelevant, provided they define the tube geometry completely.)

It is difficult to imagine that the function f should be anything but continuous and continuously increasing or decreasing with any of the variables in it. Such a function can be more or less closely approximated within a given range by a power law of some sort.

The following formula was chosen tentatively

$$\frac{\Delta R}{R} = K \frac{p}{E} \left(\frac{A}{B}\right)^m \left(\frac{R}{t}\right)^n \left(\frac{A}{t}\right)^q$$

with the unknown parameters K, m, n, and q.

Nine Bourdon tubes of different dimensions made from heat-

[7] Elliott Brothers (London) Ltd., Century Works, Lewisham, London, England.

[8] The results were published in *Instrument Practice* (London), September, 1952.

treated beryllium copper were carefully measured and tested and the results entered into the tentative equation which then could be solved for the unknown parameters.

The result is the formula

$$\frac{\Delta R}{R} = 0.05 \frac{p}{E} \left(\frac{R}{t}\right)^{1/3} \left(\frac{A}{B}\right)^{1/3} \left(\frac{A}{t}\right)^{3}$$

Two facts were surprising about this formula:

(*a*) That the exponents are round figures or reciprocals of round figures. This seems to indicate that there is, or might be, some physical significance in this formula, beyond a mere "curve fitting."

(*b*) The spread of the measured values from the values calculated from the formula is not more than 5 per cent. The conclusions to be drawn from this are that the formula represents a sound way of approximation and that tests can be made accurately enough to have a spread of this order.

The formula easily can be represented on a logarithmic graph.

Since these results were published, the formula has been used extensively in practice and, as far as the writer is aware, gives satisfactory results for Bourdon tubes of flat-oval shape for pressure ranges from about 30 to 1000 psi. For ratios A/B approximating unity, which are often used for high pressures, the formula will not hold.

There is a broad field for further experimental work to throw more light on the behavior of Bourdon tubes of various sectional shapes. The foregoing approach is of course applicable to any section shape which can be defined by a few dimensional parameters. It has the great advantage that very few tests are required to establish a practical formula which then can be applied usefully to a wide range of Bourdon tubes.

H. L. Mason.[9] The author has performed a most useful service to designers of Bourdon tubes in examining three different mathematical approaches to the problem of predicting deflections and stresses under pressure and under external moment, and reducing them to nondimensional curves which permit ready comparison. He has added a plausible approximate theory of his own, applicable to flat-ovals. A still more important contribution is his development of the theory of deflection under tip forces, which the writer has checked in detail. It uses the electrical engineers' complex-variable notation in elegant fashion, and merits publication as an Appendix to the paper.

A practical extension of the author's findings can be made if we realize that the Bourdon tube in a pressure-measuring instrument is usually called upon to drive some kind of indicating or recording device which loads the Bourdon with either moments or forces. High accuracy of measurement requires a tube which delivers a large elastic deflection for its rated internal pressure, but shows only a small tip-travel under external load. The ratio of these two deflections is a useful figure of merit for elastic design.

Consider first a pure torsional moment M to be acting on the moving tip. The resulting angular deflection is given by

$$\Delta\theta_M = \theta MR/f_3EI$$

The same Bourdon, inflated to rated pressure p_r without any M, will have an angular deflection of

$$\Delta\theta_p = \theta p_r f_1 a^4/bh^2E$$

Dividing the first into the second, letting M = unity, and taking I approximately proportional to ab^2, we obtain an accuracy criterion as

$$p_r f_1 f_3 ab(a/h)^2(a^2/Rh)$$

The larger this quantity can be made, the less effect external moment will have on pressure indication.

Again, consider a Bourdon whose free end is required to actuate a pinned link, which reacts with force F on the tip, along the line of the "larger principal flexibility" defined by the author. Angular deflection

$$\Delta\theta_F \simeq l_1/R = g_1FR^2/EIf_2$$

Proceeding as before, and noting that when 180 deg $< \theta <$ 270 deg the ratio θ/g_1 is nearly constant, we find that in this important region we can apply the same criterion

$$p_r f_1 f_2 ab(a/h)^2(a^2/Rh)$$

Making this quantity large minimizes the effect of force F.

Walter Wuest.[10] The author attributes much significance to a paper of the writer, dealing only with a special case of thin-walled Bourdon tubes, namely, with the limiting case of one cross-sectional axis being infinitely small. The solution, very complicated in general cases, becomes simple in this special case and allows one to study the effect of varying cross-sectional forms. It must be pointed out that the solution is not approximative but exact in the limit $b/a \to 0$ (a and b being the half-axes of the cross section), although no mathematical proof is given in the paper. As the sensitivity function of Bourdon tube is not greatly influenced by b/a (b/a being small), the limit solution may be regarded as an approximation of the general solution for small b/a. The author seems to have overlooked that the limit case $b/a \to 0$ of the flat-oval cross section is identical with his "ideal" cross section.

Furthermore, the writer attributes more practical usefulness to his extended work on thin and thick-walled Bourdon tubes of flat-oval cross section having arbitrary axis ratio. Of this work, done in 1942 to 1943, only recently have results been published. A complete edition is now prepared, the second part of which, dealing with high-pressure Bourdon tubes, has just been finished.[11] The theory of high-pressure Bourdon tubes (pressure range approximately 1400 to 140,000 psi) offers many interesting features quite different from that of thin-walled tubes. At very high pressures, exceeding perhaps 70,000 psi, for example, the unsymmetric plastic deformation of the cross section in the manufacturing process (bending of the tube) must be taken into account even as the elongation of the tube axis, which influences also the polar point of the tip motion. Curves for optimal construction of high-pressure Bourdon tubes are given in that report.

Author's Closure

The author agrees with Mr. Goitein that it is disappointing to find experimental points lying off an empirical curve by a factor as large as 1.5. The situation on comparing theoretical curves with one another and with experiments is even worse. On the other hand, the theories do indicate that highly curved tubes work differently from nearly straight ones. This is exemplified by the idealized section sensitivity ratio curve in Fig. 3. The straight-line portion at the left, for highly curved tubes, has a slope of $^3/_2$. The zero-slope straight-line portion at the right applies to nearly straight tubes. Note that in the transition region where the slope is changing it is not convenient to use the type of formula suggested by Mr. Goitein. To compare Goitein's

[9] General Engineer, Office of Basic Instrumentation, National Bureau of Standards, Washington, D. C. Mem. ASME.

[10] Max-Planck-Institut für Stromungsforschung, Göttingen, Germany.

[11] "Theorie der Hochdruckrohrenfeder," by Walter Wuest, preprints available from the Max-Planck-Institut für Stromungsforschung, Göttingen, Germany.

empirical formula with the idealized section straight-line portions in Fig. 3, we can give the values of K, m, n, and q. His values are 0.05, $^1/_8$, $^1/_2$, and 3, respectively. For highly curved tubes the values are radically different: 1, 3/2, 1, and 0, respectively. For nearly straight tubes the values are 2.3, 0, 1, and 3. This checks to some extent as the exponent for R/t is small and that for A/t is the same as Goitein's value.

Dr. Mason has proposed a useful figure of merit for Bourdon tubes. This can be identified as the torque that is needed to prevent the tube from deflecting. Setting $\Delta R/R = 0$ in Equation [8], we get as the moment for no deflection

$$-M = pf_1f_2Ia^4/Rbh^3$$

In case of the idealized section $I = 4\,hab^3$ and in other cases with small b/a, I is approximately proportional to hab^3. This makes the Bourdon tube torque proportional to $pf_1f_2ba^4Rh^3$. When this torque is large friction errors are minimized.

The work of Mr. Wuest on Bourdon tubes has been extremely valuable. We are looking forward to studying his "Theorie der Hochdruckrohrenfeder."

Sensitivity and Life Data on Bourdon Tubes

By H. L. MASON,[1] WASHINGTON, D. C.

Information supplied by manufacturers to the Empirical Data Subcommittee of the ASME Research Committee on Mechanical Pressure Elements is presented in tabular and graphic form. Sensitivities are compared with the theories of Wuest, Wolf, and Clark-Gilroy-Reissner. Plots of life data as a function of maximum fiber stress are shown for steel and for phosphor bronze.

Introduction

In September of 1952, the ASME Special Research Committee on Mechanical Pressure Elements authorized the appointment of a Subcommittee on Empirical Data, charged with collecting and analyzing such information on the quantitative performance of Bourdon tubes as manufacturers and users were willing to supply anonymously. The intent was to discover the current practical ranges of shape, size, material, manufacturing procedure, pressure sensitivity, compliance under force or moment, set, nonlinearity, and fatigue life. Thus a norm would be established against which mathematical theory might be checked, and from which the direction and scope of a research program might be established. The members of this committee were F. B. Jennings, General Electrical Company, C. J. Jenny, Eclipse-Pioneer Division of Bendix Corporation, I. E. Mather, Stewart-Warner Corporation, with the author as chairman.

Data Form Distributed

Data Form No. 1 diagrams the essential geometry of coil, Fig. 1, and cross section, Fig. 2, for common shapes of pressure-responsive Bourdon tubes. It provides spaces for the statement of 25 details of material, fabrication, and performance which are believed significant. To accommodate possible variations of test arrangement from firm to firm, several alternative statements of sensitivity and compliance data are provided. This form was sent to about 50 firms which were understood to be drawing Bourdon tubing, winding coils, or building these pressure-sensing devices into instruments. The letter of transmittal explained the purpose of the committee, asked for data on as many of the 25 details as the company was willing to divulge, in whatever form might be convenient, and promised that the information provided would not be coupled with the company name.

Information in varying quantities was received from ten firms, covering a total of 250 specimens. Many of these were individual tubes on which careful experimental tests had been run before they were cut apart for measurement of cross section. Other "specimens" represented design data taken from drawings or nominal specifications which factory production was expected to meet within tolerances of ±5 to 10 per cent. One manufacturer, pointing out that testing of individual specimens under RCMPE auspices would ultimately be necessary, supplied 24 mounted tubes covering a wide range of coil radii and rated pressures. Tests on these tubes, serials 19 to 42, were undertaken by O. W. Heise, Heise Bourdon Tube Company, C. J. Jenny, L. E. Smith, Bristol Co., and P. F. Erbguth, Weston Corporation, according to procedures prepared by the author.

Deflection Theory

The pressure-sensitivity data received in response to Form No. 1 covered a variety of shapes and sizes of tubes, and were expressed sometimes as an angle, sometimes as a horizontal, vertical, or radial displacement. To reduce these data to comparable figures, as required for rational analysis and design, use was made of the approximate geometry developed by Wuest (1)[2] and again by Goitein (2) and by Jennings and Duane (3). This geometry, Fig. 1 of Data Form No. 1, assumes that all cross sections of the cee-coil are identical, and that any segment of its center line is a uniform circular arc, both before and after inflation with pressure differential p.

Let the original coiled length of a segment be $R_o\phi_o$, with radius R_o and coiling angle ϕ_o, its free tip located at x_o, y_o relative to the mounting point, and its center located at 0, R_o. Upon inflation the principal change is a 1 to 5 per cent increase of radius and a corresponding decrease of angle, giving $R = R_o + \Delta R$, $\phi = \phi_o + \Delta\phi$, $R\phi = R_o\phi_o$, and $\Delta R/R \cong (-\Delta\phi)/\phi$. The tip changes position by

$$\Delta x = R \sin \phi - R_o \sin \phi_o = R_o\phi_o (\phi^{-1} \sin \phi) - R_o \sin \phi_o \quad \ldots\ldots [1a]$$

$$\Delta y = R(1 - \cos \phi) - R_o(1 - \cos \phi_o) = R_o\phi_o(\phi^{-1} - \phi^{-1} \cos \phi) - R_o(1 - \cos \phi_o) \ldots [1b]$$

Thus its total travel is

$$\sqrt{(\Delta x)^2 + (\Delta y)^2}$$

assuming that the chord approximates the arc length. The new tip position x, y is a function of the new angle ϕ, for any specified values of R_o, ϕ_o, x_o, and y_o. To find the rate at which x and y change with ϕ, take the derivatives

$$dx/d\phi = R_o\phi_o(\phi^{-1} \cos \phi - \phi^{-2} \sin \phi) \ldots\ldots [2a]$$

$$dy/d\phi = R_o\phi_o(-\phi^{-2} + \phi^{-2} \cos \phi + \phi^{-1} \sin \phi) \ldots [2b]$$

Then letting $\phi = \phi_o$, $R = R_o$, noting that $-d\phi$ is a positive quantity, and returning to finite increments

$$\Delta x/R = (-\Delta\phi/\phi)(\sin \phi - \phi \cos \phi) \ldots\ldots [3a]$$

$$\Delta y/R = (-\Delta\phi/\phi)(1 - \cos \phi - \phi \sin \phi) \ldots\ldots [3b]$$

Equations [3a] and [3b] are readily interpreted as the relative displacements in X and Y-directions, expressed as displacement factors multiplied by the relative angular displacement $(-\Delta\phi/\phi)$. Valid only for small travels, they show the direction of the tangent to the displacement arc as deflection begins. The displacement factors $(\sin \phi - \phi \cos \phi)$ and $(1 - \cos \phi - \phi \sin \phi)$ are given in Table 1, for coil angles ϕ from 180 deg to 5 turns. Thus the useful general characteristic $|\Delta\phi/\phi| = \Delta R/R$ can be computed from experimental data which may have been taken as two perpendicular displacements or as the chord $\sqrt{\Delta x^2 + \Delta y^2}$.

The theory is assumed to hold for fractional-turn cee-types, for multiturn flat spirals with average radii $R_1, R_2, \ldots R_n$, supported either at center or at periphery, and for close-wound helices.

[1] National Bureau of Standards. Mem. ASME.

Contributed by the Research Committee on Mechanical Pressure Elements and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, September 7, 1954. Paper No. 54—A-169.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

ASME Research Committee on Mechanical Pressure Elements

Data Form No. 1: Pressure-responsive Bourdon Tube

H.L.M.—Revised May 24, 1954

(1) Type: Helix or (flat) spiral or cee

(2) Material: alloy ; composition %
Young's modulus E = psi; Poisson's ratio ν =

(3) Stock before coiling: temper ; hardness
nominal grain size mm; supplied by

(4) Heat and stress treatment after coiling:

(5) Fittings soldered or brazed or welded

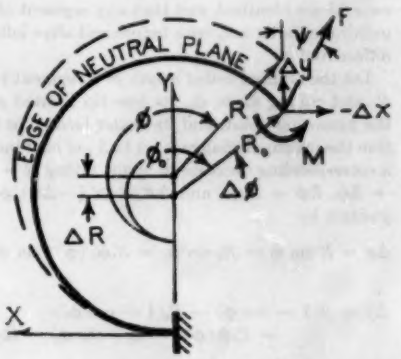


Fig. 1 Coil Geometry

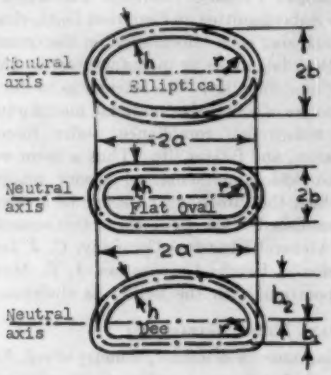


Fig. 2 Cross Sections

(6) Cross-section: elliptical or flat oval or dee Dimensions after coiling, h = a = b = b_1 = b_2 = r =
Sketch deviations if any from sections shown above.

(7) Radius and total angular length of innermost coil, R_{o1} = in., ϕ_{o1} = deg; ratio of effective (unstiffened) length of n coils to their total length = ;
for spirals supply also R_{o2}, ϕ_{o2}, ... R_{on}, ϕ_{on}

(8) Max. useful pressure diff'e p = psi; set of %, fatigue life of cycles, and nonlinearity (0 to 100% range) of % appear at temperature of deg F and this pressure difference.

(9) Pressure sensitivity, per psi: in radius, $\Delta R/R_o p$ = psi^{-1}, or in angle $-\Delta\phi/\phi_o p$ = psi^{-1}, or in deflection $\sqrt{\Delta x^2 + \Delta y^2}/R_o p$ = psi^{-1}; specify tip constraint, if any.

(10) Spring compliance, lb-in^{-1}, under pure moment M or lb^{-1} under pure force F; angle ψ = deg, at pressure p = psi: in radius $\Delta R/R_o M$ = or $\Delta R/R_o F$ = or in angle $\Delta\phi/\phi_o M$ = or $\Delta\phi/\phi_o F$ = , or in deflection $\sqrt{\Delta x^2 + \Delta y^2}/R_o M$ = or $\sqrt{\Delta x^2 + \Delta y^2}/R_o F$ = ; specify tip constraint, if any.

(11) Remarks

However, it should be noted that for (flat) spirals, Wuest limits its validity to "high-pressure springs" having $0.25 < Rh/a^2 < 1.0$ (with h and a defined by Fig. 2), and states that in "low-pressure springs" having $Rh/a^2 < 0.25$ the change of radius varies as $R^{4/3}$, the inner turns thus adding little to the total change of angle.

Pressure Sensitivity

The earliest theories of the pressure-gage element invented by Schinz, and marketed by Bourdon, as noted in the RCMPE bibliography by Van der Pyl (4), recognized that its uncoiling $|\Delta\phi|$ was proportional to coil angle ϕ, internal pressure (or vacuum) p, reciprocal modulus of elasticity E^{-1}, and a geo-

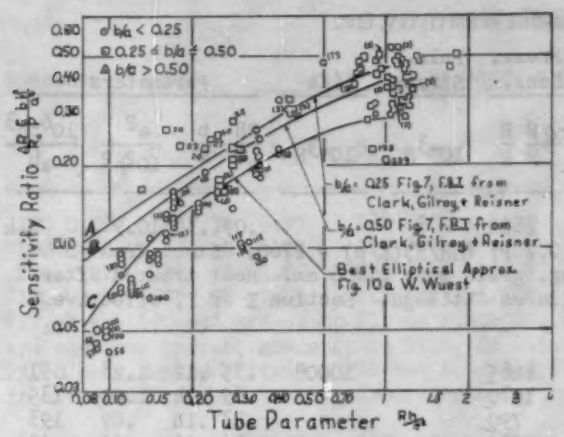

b/a < 0.25
0.25 ≤ b/a ≤ 0.50
b/a > 0.50
Sensitivity Ratio
Tube Parameter
b/a = 0.25 Fig. 7, F.R.I. from Clark, Gilroy & Reissner
b/a = 0.50 Fig. 7, F.R.I. from Clark, Gilroy & Reissner
Best Elliptical Approx. Fig. 10a W. Wuest

FIG. 3 PRESSURE SENSITIVITY OF 169 ELLIPTICAL BOURDON TUBES

metrical parameter which we shall denote by k_1. Values of k_1 have been reported here because it is a parameter recognized by all theories. Just how it should be related to the tube geometry is still an open question. In the four theories discussed by Jennings (5), it takes the form f_1a^4/bh^3, and he plots

$$f_1 = (\Delta R/R)(E/p)(bh^3/a^4)$$

as a function of Rh/a^2.

Table 2 lists the information available on pressure sensitivity and life of our 250 specimens. Values followed by an asterisk are known to be only nominal, and the shapes designated are probably nominal. The column headings represent the following:

1 Specimen number, e.g., 160 S 002: The first figure is the manufacturer's pressure rating p in psi, the letter is the coil type (C = cee, S = flat spiral, H = helical) and the last figure is a serial number, assigned by the subcommittee.

2 Stock description: Wall thickness h in mils followed by an abbreviation for the material; BC = beryllium copper, Bz = bronze, CM = chrome-molybdenum steel, CRS = corrosion-resistant steel, CS = carbon steel, PB = phosphor bronze, S = steel, TM = trumpet metal.

3 Coil description: Radius R_0 in inches, angle ϕ_0 in degrees; no allowance has been made for end fittings.

4 Cross-section description, Fig. 2: Major semiaxis a at mid-wall, mils; minor semiaxis b at mid-wall, mils; approximate shape E = elliptical, F = flat oval, D = dee.

5 Pressure sensitivity $k_1 = (-\Delta\phi/\phi)(E/p)$, dimensionless; modulus E, unless otherwise noted, is taken from Giacobbe and Bounds (6); angle changes computed from travel as Δx, Δy, or $\sqrt{\Delta x^2 + \Delta y^2}$ are so noted.

6 Life, $10^{-3}N$ cycles; from 0 to rated pressure p unless otherwise noted, and to fracture, unless otherwise noted.

7 Pulsation stress (maximum fiber stress) developed on life test, $10^{-3}S$ psi.

8 Parameters Rh/a^2, b/a, $a^2/10^3h^2$, $10^6bh^3/a^4$, all dimensionless; useful for comparison with various theories (5).

At the request of the RCMPE Subcommittee on Theory, the sensitivity data of Table 2 have been plotted with ordinate log $(\Delta R/R_0)(E/p)(bh^3/a^4)$ and abscissa log Rh/a^2. Fig. 3 shows data on specimens of elliptical cross section in comparison with three curves from Jennings (5); A and B based on the Clark-Gilroy-Reissner theory (7) and C based on the Wuest theory (8). Points are marked with a specimen number if they differ widely from theory; those representing several specimens are followed by a number in parentheses. In the region $0.09 < Rh/a^2 < 0.2$, most of the points for $b/a < 0.25$ (circled) cluster around curve C, the trend being somewhat steeper. In the region $0.2 < Rh/a^2 < 0.5$, the points for $b/a < 0.25$ tend to fall between curves C and A, the trend being steeper than either curve. Above $Rh/a^2 = 0.5$, the circled points are too few for one to generalize. In the region $0.18 < Rh/a^2 < 1.8$, the points for $0.25 \leq b/a \leq 0.50$ (squared) tend to follow curve B, although the scatter is about ±25 per cent. Specimen No. 1 is a maverick, with a sensitivity ratio of 0.022 at $Rh/a^2 = 0.09$.

TABLE 1 DISPLACEMENT FACTORS

Coil angle	$(\sin\phi - \phi\cos\phi)$	$(1 - \cos\phi - \phi\sin\phi)$
90 deg	+1	−0.57
180	+3.14	+2
210	+2.67	+3.70
240	+1.22	+5.12
270	−1	+5.71
300	−3.49	+5.03
330	−5.48	+3.01
360	−6.28	0
1.5 turns	+9.42	+2
2.0	−12.57	0
2.5	+15.71	+2
3.0	−18.85	0
3.5	+21.99	+2
4.0	−25.13	0
4.5	+28.27	+2
5.0	−31.42	0

Fig. 6 shows data on specimens of flat-oval and dee cross section, in comparison with five curves from Jennings (5), A and B based on Wuest theory (8), C and D based on Wolf theory (9), and E based on Jennings approximation. For points with $b/a > 0.3$ (triangular), E is a better approximation than either B or D. For $0.3 > b/a > 0.1$ (squared points), B is a slightly better approximation than the other curves. The helical group No. 7–12 lie roughly parallel to B and E, but show only about half the theoretical sensitivity, perhaps due to the tip restraint.

LIFE

For the pulsation stress developed on fatigue test, measured values are available for only a few; these are based on stress-coat readings at the exterior ends of Fig. 2(*a*). The other values reported were computed from

$$S = 1.991\, p(a/h)^2(1 - b^2/a^2)(a/b)(Rh/a^2)^{1/4}$$

with Poisson ratio $\nu = 1/3$, from Jennings and Duane (3), based on Clark-Gilroy-Reissner (7).

Since it is well known that ferrous and nonferrous materials behave differently under repetitive stressing, separate fatigue curves were plotted for these elliptical groups. Fig. 4 shows S versus log N for steel specimens. The band of points is about 80,000 psi wide, but shows a tendency to level off at an average of about 110,000 psi.

Fig. 5 shows S versus log N for phosphor-bronze specimens. For them the fatigue life is about 10,000 cycles at 100,000 psi, 100,000 cycles at 75,000 psi, and 1,000,000 cycles at 50,000 psi. However, the range of variation is of the order of ±40,000 psi, and there is wide variation between nominally identical specimens.

CONCLUSION

It is hoped that both RCMPE members and the firms which so kindly supplied data will study this report in the light of their own experience with Bourdon tubes, and proffer discussions which shed light on the anomalies noted herein. It is urged that RCMPE take steps to secure more precise data in the region $0.05 < Rh/a^2 < 2$, and to fill in the unexplored regions above and below this. Data on the compliance of Bourdons to forces or moments is on hand for about 30 specimens, and work is under way on reducing this information to comparable figures useful to the designer.

TABLE 2 GEOMETRY, PRESSURE SENSITIVITY, LIFE

Specimen		Stock		Coil		Section		Press. Sens.	Puls. Stress	Life	Parameters			
p	#	h	Mtl.	R_o	ϕ_o	a	b	$\frac{-\Delta\phi}{\phi}\frac{E}{p}$	$10^{-3}S$	$10^{-3}N$	$\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
100	C 001	25	PB	2.55	255	825	x250 E	2620	3		.09	.30	10.9	8.4

Serial 001 is phosphor bronze of 95 Cu, 5 Sn, 0.2 P; $(\Delta R/R)(E/p)$ = 2760. Stock from Mackenzie Walton, half hard, Rockwell B-78, avg. grain size 0.886 mm. Heat treated after coiling, 500F for 1 hr, to Rockwell B-97. Soldered fittings. Section E or F, effective ϕ_o=248. Set < 0.05% at 70F.

160	S 002	18	BC		See	200	x 25 F	1485		1000[a]	.35	.12	1.23	91
230	S 003	20	BC		note	200	x 27 F	1030		"	.39	.14	1.00	135
300	S 004	22	BC		below	200	x 29 F	790		"	.43	.14	.82	193
370	S 005	24	BC			200	x 31 F	640		"	.46	.16	.69	268
510	S 006	28	BC			200	x 35 F	465		"	.54	.18	.51	480

Serials 2 to 6 are beryllium copper of 1.0-2.15 Be, 0.35 Co, bal. Cu; E=19x10^6, ν=0.3; 4.5 turn spirals of 1600°, with uniformly spaced radii, innermost 0.5" to outermost 1.0", fixed at outer end, 20° rotation at inner end, mean radius R_o taken as 0.77". Stock is solution annealed to Rockwell 30T46-60 before coiling; heattreated at 650 ± 5 F for 60 ± 5 min. after coiling. Fittings silver soldered.
Note a: For 1% set; 0% nonlinearity; 80°F.

90	H 007	11.5*TB		0.95*	1440	386*x	33*F	2660	39	200[b]	.07	.09	1.12	2.3
120	H 008	12.5*TB		0.95*	1440	387*x	34*F	2000	45	"	.08	.09	.96	3.0
270	H 009	18*	TB	0.96*	1440	393*x	43*F	886	52	"	.11	.11	.48	10.5
275	H 010	21*	TB	1.57*	1440	383*x	63*F	871	40	"	.22	.16	.33	27.1
500	H 011	16*	TB	1.05*	3240	244*x	126*F	210	27	"	.24	.52	.23	146
200	H 012	21*	TB	0.97*	810	635*x	42*F	2100	68	"	.05	.07	.91	2.4

Serials 7 to 12 are trumpet brass of 81-82 Cu, 0.75-1.5 Sn, 0.05 max P, 0.05 max Fe, bal Zn, E=15x10^6, ν=0.33. Tip is constrained to radius R_o. Stock from American Brass Co. and Mackenzie Walton, cold drawn, hardness under .020" 15T-82 to 89, 30T-62 to 72, grain size .010-.030 mm. Heat treated at 480 F for 1 hr, furnace cooled with nitrogen. Fittings soldered. Nonlinearity 1/2% on #11, 1% on #7 to 10, 1 1/2% on #12.
Note b: For set of 1/2%, room temp.

30	C 019	10*	TB	0.97	233	273	x 87 E	10700	-	-	.13	.32	7.4	15.8
30	C 020	10.5*TB		1.19	235	275	x 84 E	9800	-	-	.16	.31	6.9	18.3
30	C 021	12*	PB	1.62	228	407	x108 F	8500	-	-	.12	.27	11.4	9.5
30	C 022	15*	PB	2.06	241	407	x120 F	7200	-	-	.19	.29	7.4	14.7

#19,20 are trumpet brass; Cu 81, Sn 1, Zn 18; E=15x10^6 psi, G=6x10^6 psi, ν=0.25.
#21,22 are phosphor bronze, grade A; Cu 95.5, Sn 4.5; E=16x10^6 psi, G=6.25x10^6 psi, ν=0.28.
#19,21,22 have brazed fittings, #20 has soldered fittings. Not pulsed. $\Delta\phi/\phi$ based on chordal travel; data averaged from two independent testing laboratories.

60	C 023	14*	TB	0.97	233	275	x 88 E	5000	-	-	.18	.32	3.8	42.2
60	C 024	14.5*TB		1.17	237	276	x 87 E	5000	-	-	.22	.32	3.6	45.7
60	C 025	17*	PB	1.62	228	408	x119 F	4000	-	-	.17	.29	5.7	21.2
60	C 026	21*	PB	2.06	241	409	x125 F	3800	-	-	.26	.31	3.8	41.4

#23,24 are trumpet brass; Cu 81, Sn 1, Zn 18; E=15x10^6 psi, G=6x10^6 psi, ν=0.25. #25,26 are phosphor bronze, grade A; Cu 95.5, Sn 4.5; E=16x10^6 psi, G=6.25x10^6 psi, ν=0.28.
#23,25,26 have brazed fittings, #24 has soldered fittings. Not pulsed. $\Delta\phi/\phi$ based on chord; data averaged from two independent testing laboratories.

TABLE 2 (continued)

Specimen	Stock		Coil		Section		Press. Sens.	Puls. Stress	Life	Parameters			
p #	h	Mtl.	R_o	ϕ_o	a	b	$\frac{-\Delta\phi}{\phi}\frac{E}{p}$	$10^{-3}S$	$10^{-3}N$	$\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
100 C 027	18*	TB	0.97	233	276 x 84	E	3000	-	-	.23	.30	2.3	84.5
100 C 028	18*	TB	1.17	237	274 x 93	E	3000	-	-	.28	.34	2.3	96.6
100 C 029	21*	PB	1.62	228	412 x123	F	2400	-	-	.20	.30	3.8	39.3
100 C 030	26*	PB	2.06	241	413 x127	F	2300	-	-	.31	.31	2.5	76.6
200 C 031	21*	TB	1.00	232	254 x 48	D	1410			.32	.19	1.5	105
200 C 032	23*	PB	1.18	234	255 x 48	D	1390			.42	.19	1.2	136
200 C 033	30*	PB	1.62	221	412 x129	D	1400			.29	.31	1.9	122
200 C 034	35*	PB	2.05	242	416 x123	D	1240			.42	.30	1.4	176
600 C 035	34*	TB	1.00	230	264 x 53	D	506			.49	.20	.6	431
600 C 036	36*	PB	1.19	234	262 x 56	D	430			.62	.21	.5	550
600 C 037	43*	PB	1.60	226	340 x 82	D	512			.60	.24	.6	588
600 C 038	48*	PB	2.00	244	344 x 85	D	408			.82	.25	.5	678
1000 C 039	41*	TB	1.02	230	264 x 56	D	270			.60	.21	.4	796
1000 C 040	42*	PB	1.20	235	266 x 54	D	304			.71	.20	.4	764
1000 C 041	54*	PB	1.62	226	352 x 88	D	277			.70	.25	.4	912
1000 C 042	58*	PB	2.02	244	350 x 87	D	246			.96	.25	.4	1150

#27,28 are trumpet brass; Cu 81, Sn 1, Zn 18; $E=15\times10^6$ psi, $G=6\times10^6$ psi, $\nu=0.25$. #29,30 are phosphor bronze, grade A; Cu 95.5, Sn 4.5; $E=16\times10^6$ psi, $G=6.25\times10^6$ psi, $\nu=0.28$. #27, 29,30 have brazed fittings, #28 has soldered fittings. $\Delta\phi/\phi$ based on chord; data averaged from two independent testing laboratories.

Serials 31 to 42 made from elliptical tubing, flattened to D-shape in coiling; neutral axis assumed central. Data are averaged from two independent testing laboratories; $\Delta\phi/\phi$ based on chord. Rated pressure applied just before deflection test, #31-42. Serials 31,33-42 are silver soldered or brazed; #32 is soft soldered. Serials 31, 35, 39 are trumpet brass; Cu 81, Sn 1, Zn 18; $E=15\times10^6$ psi, $G=6\times10^6$ psi, $\nu=0.25$. Serials 32-34, 36-38, 40-42 are phosphor bronze grade A; Cu 95.5, Sn 4.5; $E=16\times10^6$ psi, $G=6.25\times10^6$ psi, $\nu=0.28$.

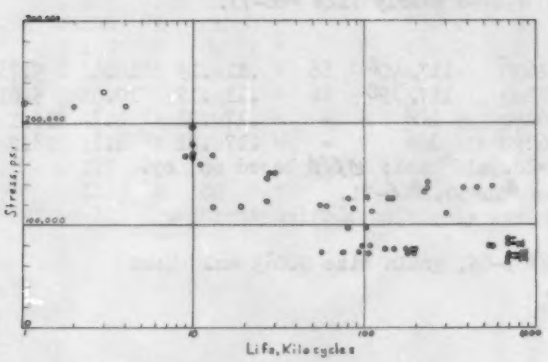


FIG. 4 FATIGUE OF STEEL BOURDON TUBES

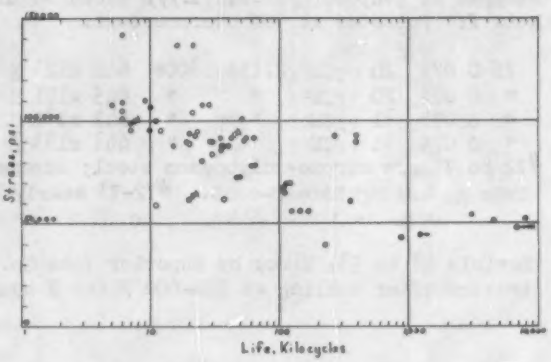


FIG. 5 FATIGUE OF PHOSPHOR-BRONZE BOURDON TUBES

TABLE 2 (*continued*)

Specimen	Stock		Coil		Section	Press. Sens.	Puls. Stress	Life	Parameters			
p #	h	Mtl.	R_o	ϕ_o	a b	$-\frac{\Delta\phi}{\phi}\frac{E}{p}$	$10^{-3}S$	$10^{-3}N$	$\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
500 C 043	14	BC	0.71	245	134 x 42 F	930	-	-	.56	.33	..96	356
400 C 044	14	BC	0.71	238	140 x 31 F	1310	-	-	.51	.22	1.00	221
200 C 045	14	BC	0.72	250	146 x 18 F	2100	-	-	.48	.12	1.08	108
2000 C 046	22	BC	0.72	245	130 x 36 F	230	-	-	.93	.28	.35	1340
1400 C 047	22	BC	0.71	250	135 x 23 F	360	-	-	.85	.17	.38	738
800 C 048	22	BC	0.75	248	141 x F	560	-	-	.83	.08	.41	
5000 C 049	32	BC	0.73	248	127 x F	91	-	-	1.44	.17	.16	
5000 C 050	32	BC	0.72	250	120 x 35 F	57	-	-	1.59	.29	.14	5500
5000 C 051	32	BC	0.74	250	115 x 46 F	39	-	-	1.78	.40	.13	8500
4000 C 052	30	BC	0.80	300	124 x F	97	-	-	1.56	.24	.17	
4000 C 053	30	BC	0.82	300	117 x 48 F	53	-	-	1.80	.41	.15	6840
#43 to 53 are beryllium copper; assume E=19.5x10^6 psi.												
50 C 054	20	CM	2.25*	300*	648 x123 E	12300	80	535	.11	.19	10	5.59
" C 055	-	CM	"	"	648 x126 E	13400	-	167	-	.19	-	
" C 056	20	CM	"	"	658 x150 E	6500	112	301[f]	.10	.23	11	6.35
" C 057	18	CM	"	"	670 x156 E	9600	136	233[f]	.09	.23	14	4.51
" C 058	18	CM	"	"	661 x152 E	9900	-	-	.09	.23	14	4.61
" C 059	18	CM	"	"	660 x154 E	10400	133	225[f]	.09	.23	13	4.74
" C 060	14	CM	"	"	363 x119 E	14000	79	551[f]	.24	.33	7	18.7
" C 061	12.5	CM	"	"	420 x100 E	17100	87	684	.16	.24	11	6.26
" C 062	13	CM	"	"	418 x102 E	17100	82	684	.17	.24	10	7.35
" C 063	13	CM	"	"	419 x101 E	16600	82	684	.17	.24	10	7.21
" C 064	-	CM	"	"	418 x100 E	17300	-	684	-	.24	-	
" C 065	13	CM	"	"	418 x100 E	17600	82	684	.17	.24	10	7.20
" C 066	20	CM	"	"	642 x126 E	14900	76	131	.11	.20	10	5.92
" C 067	20	CM	"	"	642 x124 E	15300	76	148	.11	.19	10	5.84
" C 068	20	CM	"	"	642 x128 E	15700	76	196	.11	.20	10	6.03
" C 069	20	CM	"	"	642 x126 E	15100	-	-	.11	.20	10	5.92
" C 070	20	CM	"	"	642 x125 E	15100	76	190	.11	.19	10	5.89
" C 071	20	CM	"	"	640 x126 E	15500	76	175	.11	.20	10	5.95
#54 to 71 are chrome-molybdenum steel; assume E=28.5x10^6 psi. Life test on #61 to 65 stopped at 684,000 cycles. $\Delta\phi/\phi$ based on Δy. #54-56 nearly like #66-73.												
Note f: Pulsated at 50% overpressure.												
75 C 072	20	CM	2.25*	300*	644 x123 E	12600	117, 40[g]	86	.11	.19	10.4	5.73
" C 073	20	CM	"	"	645 x121 E	12200	117, 39[g]	86	.11	.19	10.4	5.61
" C 074	31	CM	"	"	642 x132 E	6290	40[g]	-	.17	.21	4.3	23.1
" C 075	31	CM	"	"	643 x133 E	6290	38[g]	-	.17	.21	4.3	23.2
#72 to 75 are chrome-molybdenum steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.												
Note g: Using Stress-coat. #72-73 nearly like #54-56, #66-71.												

Serials 43 to 53: Stock by Superior Tube Co., R30T-65, grain size 0.035 mm. Heat treated after coiling at 550-600 F for 2 hours.

TABLE 2 (continued)

Specimen		Stock		Coil		Section		Press. Sens.	Puls. Stress	Life	Parameters			
p	#	h	Mtl.	R_o	ϕ_o	a	b	$\frac{-\Delta\phi}{\phi}\frac{E}{p}$	$10^{-3}S$	$10^{-3}N$	$\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
100	C 076	23	CM	2.25*	300*	682	x111 E	12800	142	235	.11	.16	8.8	6.23
"	C 077	24	CM	"	"	665	x131 E	8390	118, 118	60	.12	.20	7.5	9.21
"	C 078	23	CM	"	"	661	x132 E	8700	127, 118	144	.12	.20	8.3	8.34
"	C 079	23	CM	"	"	661	x131 E	8390	127	81	.12	.20	8.3	8.29
"	C 080	23	CM	"	"	661	x131 E	8960	127	103	.12	.20	8.3	8.29
"	C 081	23	CM	"	"	667	x141 E	7820	-	-	.12	.21	8.4	8.62
"	C 082	23	CM	"	"	661	x144 E	7690	-	-	.12	.22	8.3	9.12
"	C 083	23	CM	"	"	666	x144 E	7790	-	-	.12	.22	8.4	8.80
"	C 084	23	CM	"	"	665	x144 E	7820	-	-	.12	.22	8.3	8.94
"	C 085	20	CM	"	"	362	x114 E	4420	-	-	.34	.32	3.3	53.5
"	C 086	20	CM	"	"	363	x114 E	4400	-	-	.34	.31	3.3	53.8
"	C 087	22	CM	"	"	408	x103 E	5800	64	684	.30	.25	3.4	39.4
"	C 088	21	CM	"	"	406	x109 E	6190	68	684	.28	.27	3.8	37.2
"	C 089	21	CM	"	"	409	x102 E	5850	71	684	.28	.25	3.8	33.8
"	C 090	21	CM	"	"	406	x104 E	5900	69	684	.28	.26	3.8	35.6
"	C 091	21	CM	"	"	410	x101 E	6190	71	684	.28	.25	3.8	33.2
"	C 092	19	CM	"	"	408	x112 E	6290	-	-	.25	.27	4.6	28.0
"	C 093	19	CM	"	"	407	x113 E	6010	-	-	.25	.28	4.6	28.4
"	C 094	19	CM	"	"	407	x112 E	6140	-	-	.25	.27	4.6	28.0
"	C 095	19	CM	"	"	408	x113 E	6010	-	-	.25	.28	4.6	28.2
"	C 096	19	CM	"	"	409	x110 E	5940	-	-	.26	.27	4.6	26.8
"	C 097	19	CM	"	"	415	x101 E	6690	-	-	.25	.27		25.5
"	C 098	19	CM	"	"	414	x101 E	6690	-	-	.25	.24		23.6
"	C 099	19	CM	"	"	414	x101 E	7060	-	-	.25	.24		23.6
"	C 100	19	CM	"	"	411	x107 E	6430	-	-	.25	.26		25.2
"	C 101	19	CM	"	"	410	x107 E	6400	-	-	.25	.26		25.2
"	C 102	19	CM	"	"	410	x107 E	7000	-	-	.25	.26		25.2

#76 to 102 are X4130 steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy. Life test on #87 to 91 stopped at 684,000 cycles.

100	C 103	31	CM	2.25*	300*	636	x126 E	6350	72	105	.17	.20	4.2	23
"	C 104	31	CM	"	"	635	x129 E	6010	72	196	.17	.20	4.2	24
"	C 105	31	CM	"	"	638	x126 E	6350	72	180	.17	.20	4.2	22
"	C 106	31	CM	"	"	638	x125 E	6350	-	-	.17	.20	4.2	22
"	C 107	31	CM	"	"	636	x125 E	6350	-	-	.17	.20	4.2	23
"	C 108	31	CM	"	"	636	x125 E	6350	-	-	.17	.20	4.2	23
"	C 109	31	CM	"	"	637	x121 E	6990	73	56	.17	.19	4.2	22
"	C 110	31	CM	"	"	638	x121 E	7190	73	75	.17	.19	4.2	22
"	C 111	31	CM	"	"	638	x119 E	7010	74	93	.17	.19	4.2	21
"	C 112	30	CM	"	"	640	x120 E	7010	79	107	.16	.19	4.6	19
"	C 113	30	CM	"	"	641	x120 E	6810	79	97	.16	.19	4.6	19

#103 to 113 are chrome-molybdenum steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.

TABLE 2 (continued)

Specimen		Stock		Coil		Section	Press. Sens.	Puls. Stress	Life	Parameters			
p	#	h	Mtl.	R_o	ϕ_o	a b	$-\frac{\Delta\phi}{\phi}\frac{E}{p}$	$10^{-3}S$	$10^{-3}N$	$\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
250	C 114	40	CM	2.25*	300*	642 x127 E	3980	120,37E	55	.22	.20	2.6	48
"	C 115	40	CM	"	"	640 x134 E	3720	118,35E	19	.22	.21	2.6	51
"	C 116	56	CM	"	"	676 x115 E	1480	-	-	.28	.17	1.5	94
"	C 117	56	CM	"	"	644 x130 E	1400	137	379	.30	.20	1.3	133
"	C 118	56	CM	"	"	642 x132 E	1430	138	556	.31	.21	1.3	136
"	C 119	56	CM	"	"	643 x131 E	1370	137	453	.30	.20	1.3	134
"	C 120	59	CM	"	"	641 x129 E	1430	126	138	.32	.20	1.2	155
"	C 121	27	CM	"	"	366 x 98 E	3250	98	101	.45	.27	1.8	107
"	C 122	27	CM	"	"	365 x 98 E	3020	-	-	.46	.27	1.8	107
"	C 123	27	CM	"	"	365 x 99 E	3020	98	80	.46	.27	1.8	108
#114 to 123 are chrome-molybdenum steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.													
#117-120 pulsated at 500 psi.													
500	C 124	63	CM	2.25*	300*	632 x 90 E	1770	36E	7	.35	.14	1.0	142
"	C 125	61	CM	"	"	635 x 93 E	1850	117,37E	13	.34	.15	1.0	146
"	C 126	73	CM	"	"	631 x124 E	740	84	-	.41	.22	0.8	296
"	C 127	73	CM	"	"	631 x124 E	770	-	-	.41	.22	0.8	296
"	C 128	73	CM	"	"	631 x124 E	740	-	-	.41	.22	0.8	296
"	C 129	73	CM	"	"	631 x124 E	740	-	-	.41	.22	0.8	296
"	C 130	39	CM	"	"	362 x 91 E	1250	113	110	.67	.25	0.9	314
"	C 131	39	CM	"	"	362 x 91 E	1230	-	-	.67	.25	0.9	314
"	C 132	39	CM	"	"	362 x 91 E	1250	-	-	.67	.25	0.9	314
#124 to 132 are chrome-molydbenum steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.													
1000	C 133	42,48	CM	2.25*	300*	354 x104 E	740	151	30	.81	.29	.62	600
"	C 134	42,49	CM	"	"	354 x103 E	770	146	27	.83	.29	.59	635
"	C 135	41,47	CM	"	"	355 x104 E	800	124	27	.79	.29	.51	547
"	C 136	40,49	CM	"	"	354 x104 E	800	151	28	.81	.29	.62	600
#133 to 136 are chrome-molybdenum steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy. Outer wall thinner than inner wall.													
1500	C 137	51	CM	2.25*	300*	348 x116 E	480	168	9	.95	.33	.47	1040
"	C 138	51	CM	"	"	349 x118 E	480	166	10	.94	.34	.47	1060
"	C 139	50	CM	"	"	347 x118 E	430	169	10	.93	.34	.48	1000
"	C 140	52	CM	"	"	347 x118 E	400	160	11	.97	.34	.45	1130
"	C 141	50	CM	"	"	348 x120 E	480	169	13	.93	.34	.48	1010
"	C 142	47	CM	"	"	345 x 86 E	510	221	1	.89	.25	.54	640
"	C 143	48	CM	"	"	345 x 84 E	510	216	2	.91	.24	.52	660
"	C 144	46	CM	"	"	345 x 87 E	510	231	3	.86	.25	.57	570
"	C 145	47	CM	"	"	340 x 83 E	480	216	4	.92	.24	.52	640
#137 to 145 are chrome-molybdenum steel; assume E=28.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.													

TABLE 2 (*continued*)

Specimen p	#	Stock h	Mtl.	Coil R_o	ø	Section a	b	Press. Sens. $\frac{-\Delta\phi}{\phi}\frac{E}{p}$	Puls. Stress $10^{-3}S$	Life $10^{-3}N$	Parameters $\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
25	C 146	18	PB	2.25*	300*	645	x108 E	14500	48	684	.097	.17	12.8	3.63
"	C 147	18	PB	"	"	645	x106 E	14600	48	684	.097	.16	12.8	3.56
"	C 148	18	PB	"	"	646	x108 E	14000	48	684	.097	.17	13.0	3.59
"	C 149	18	PB	"	"	647	x108 E	14300	48	684	.097	.17	13.0	3.59
"	C 150	18	PB	"	"	646	x105 E	13500	-	-	.097	.16	13.0	3.50
"	C 151	18	PB	"	"	645	x107 E	14300	-	-	.097	.17	12.8	3.60
#146 to 151 are phosphor bronze; assume E=15.5x10^6 psi; $\Delta\phi/\phi$ based on Δy. Life test stoped at 684,000 cycles.														
50	C 152	26	PB	2.25*	300*	659	x113 E	9100	54	3040	.14	.17	6.4	10.5
"	C 153	27	PB	"	"	658	x114 E	9100	51	4565	.14	.18	6.0	11.9
"	C 154	27	PB	"	"	643	x106 E	7500	-	-	.15	.16	6.2	12.2
"	C 155	27	PB	"	"	642	x107 E	7200	-	-	.15	.17	6.2	12.3
"	C 156	27	PB	"	"	658	x107 E	7200	52	799	.14	.16	6.0	11.2
"	C 157	32	PB	"	"	640	x103 E	7200	39	224	.18	.16	4.0	20.1
"	C 158	29	PB	"	"	642	x104 E	7700	44	1224	.16	.16	4.9	14.8
"	C 159	30	PB	"	"	642	x103 E	7200	43	870	.16	.16	4.6	16.3
#152 to 159 are phosphor bronze; assume E=15.5x10^6 psi; $\Delta\phi/\phi$ based on Δy. Life test on #158 stopped at 1,224,000 cycles.														
100	C 160	27	PB	2.25*	300*	648	x134 E	4400	93	265	.14	.21	5.8	15
"	C 161	28	PB	"	"	652	x135 E	5300	90	386	.15	.21	5.5	18
"	C 162	27	PB	"	"	652	x138 E	5300	93	391	.14	.21	5.9	17
"	C 163	37	PB	"	"	636	x 96 E	4500	61	352	.21	.15	3.0	30
"	C 164	38	PB	"	"	635	x 99 E	4200	56	169	.21	.16	2.8	33
"	C 165	38	PB	"	"	635	x102 E	4200	56	145	.21	.16	2.8	34
"	C 166	38	PB	"	"	635	x 99 E	4300	56	125	.21	.16	2.8	33
#160 to 166 are phosphor bronze; assume E=15.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.														
250	C 167	60	PB	2.25*	300*	630	x 93 E	1800	-	-	.34	.15	1.1	127
"	C 168	59	PB	"	"	629	x 94 E	1500	69	113	.34	.15	1.1	125
"	C 169	61	PB	"	"	628	x 93 E	1500	65	110	.35	.15	1.1	135
"	C 170	59	PB	"	"	631	x 94 E	1500	70	119	.35	.15	1.1	121
"	C 171	62	PB	"	"	628	x 93 E	1500	62	37	.35	.15	1.0	145
"	C 172	60	PB	"	"	630	x 93 E	1500	68	107	.34	.15	1.1	127
#167 to 172 are phosphor bronze; assume E=15.5x10^6 psi; $\Delta\phi/\phi$ based on Δy.														
500	C 173	99	PB	2.25*	300*	616	x 85 E	820	59	11	.59	.14	.39	570
"	C 174	-	PB	"	"	-	x - E	850	-	-	-	-	-	-
"	C 175	-	PB	"	"	-	x - E	850	-	-	-	-	-	-
"	C 176	72	PB	"	"	629	x100 E	820	96	12	.41	.16	.76	420
"	C 177	73	PB	"	"	630	x101 E	-	93	22	.41	.16	.74	400
"	C 178	76	PB	"	"	624	x 85 E	800	92	42	.44	.14	.67	400
"	C 179	77	PB	"	"	623	x 85 E	800	92	55	.45	.14	.66	390
"	C 180	76	PB	"	"	624	x 86 E	800	92	63	.44	.14	.67	400
"	C 181	76	PB	"	"	624	x 86 E	790	92	43	.44	.14	.67	400
"	C 182	76	PB	"	"	624	x 86 E	790	91	53	.44	.14	.67	390
Phosphor bronze, E=15.5x10^{6}psi; $\Delta\phi/\phi$ based on Δy.														

TABLE 2 *(continued)*

Specimen p #	Stock h	Mtl.	Coil R_o	ϕ_o	Section a x b	Press. Sens. $-\frac{\Delta\phi}{\phi}\frac{E}{p}$	Puls. Stress $10^{-3}S$	Life $10^{-3}N$	Parameters $\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
500 - 183	41.5	-	-	-	338 x105 E	-	-	13	-	.31	.65	580
" - 184	48	-	-	-	335 x101 E	-	-	32	-	.30	.49	890
" - 185	47	-	-	-	334 x104 E	-	-	46	-	.31	.51	870
" - 186	48	-	-	-	333 x105 E	-	-	88	-	.32	.48	960
" - 187	48	-	-	-	335 x105 E	-	-	120	-	.31	.49	920
" - 188	48	-	-	-	- x101 E	-	-	46	-	-	-	-
" - 189	48	-	-	-	- x102 E	-	-	52	-	-	-	-
" - 190	48	-	-	-	- x105 E	-	-	52	-	-	-	-
" - 191	49	-	-	-	339 x 91 E	-	-	49	-	.27	.43	810
" - 192	41	-	-	-	342 x 94 E	-	-	14	-	.27	.69	470
#183,188,189 and 190 are "double tubes".												
1000 C 193	57.5	PB	2.25*	300*	380 x131 E	190	101	58	.90	.34	.44	1190
" C 194	66	PB	"	"	375 x128 E	190	78	61	1.06	.34	.32	1900
" C 195	68	PB	"	"	376 x126 E	190	78	63	1.11	.34	.31	1970
" C 196	55	PB	"	"	337 x 95 E	390	101	6	1.09	.28	.37	1230
" C 197	56	PB	"	"	335 x 97 E	390	99	7	1.12	.29	.36	1380
" C 198	56	PB	"	"	335 x 98 E	390	99	7	1.12	.29	.36	1390
" C 199	57	PB	"	"	336 x 91 E	390	96	10	1.04	.27	.35	1330
" C 200	54	PB	"	"	336 x 96 E	390	107	10	1.08	.29	.39	1190
" C 201	56	PB	"	"	336 x 94 E	390	100	11	1.12	.28	.36	1270
" C 202	59	PB	"	"	333 x 94 E	-	90	19	1.20	.28	.32	1570
" C 203	57.5	PB	"	"	332 x 93 E	-	94	20	1.17	.28	.33	1470
" C 204	58	PB	"	"	331 x 93 E	280	90	30	1.19	.28	.32	1480
" C 205	60	PB	"	"	329 x 93 E	280	84	30	1.17	.28	.30	1660
" C 206	59	PB	"	"	329 x 93 E	170	86	35	1.16	.28	.31	1660
" C 207	59	PB	"	"	330 x 93 E	280	88	37	1.22	.28	.31	1660
" C 208	51	PB	"	"	333 x108 E	260	108	5	1.03	.32	.43	1160
" C 209	51	PB	"	"	336 x106 E	190	109	6	1.02	.32	.43	1090
" C 210	50	PB	"	"	334 x107 E	360	110	9	1.01	.32	.44	1070
" C 211	51	PB	"	"	335 x108 E	340	108	25	1.02	.32	.43	1150
" C 212	51	PB	"	"	334 x108 E	360	108	28	1.03	.32	.43	1150
" C 213	48	PB	"	"	288 x 80 E	390	105	6	1.31	.28	.36	1260
" C 214	40	PB	"	"	292 x 86 E	390	142	6	1.06	.29	.53	760
" C 215	40	PB	"	"	291 x 90 E	410	137	16	1.06	.31	.53	810
" C 216	47	PB	"	"	283 x 82 E	360	104	21	1.32	.29	.36	1320
" C 217	40	PB	"	"	291 x 89 E	390	137	21	1.06	.31	.53	800
" C 218	55	PB	"	"	331 x101 E	230	96	7	1.13	.30	.36	1390
" C 219	56	PB	"	"	330 x102 E	230	95	23	1.16	.30	.35	1430
" C 220	57	PB	"	"	330 x100 E	190	90	29	1.18	.30	.33	1590
" C 221	58	PB	"	"	329 x100 E	210	88	36	1.21	.30	.32	1690
" C 222	58	PB	"	"	329 x100 E	210	88	40	1.21	.30	.32	1690
" C 223	60	PB	"	"	332 x 96 E	210	88	15	1.22	.29	.31	1700
" C 224	59	PB	"	"	332 x 98 E	210	89	20	1.20	.29	.32	1640
" C 225	59	PB	"	"	331 x 98 E	210	92	35	1.21	.30	.33	1560
" C 226	64	PB	"	"	367 x125 E	180	81	25	1.07	.34	.33	2280
" C 227	68	PB	"	"	365 x125 E	150	73	49	1.15	.34	.29	2280
" C 228	68	PB	"	"	365 x125 E	170	73	47	1.15	.34	.29	2860
" C 229	73	PB	"	"	361 x124 E	130	62	20	1.26	.34	.24	2860
" C 230	74	PB	"	"	363 x122 E	150	62	20	1.26	.34	.24	2800

#193 to 230 are phosphor bronze; assume E=15.5x10^6 psi; $\Delta\phi/\phi$ based on Δy; #208 to 212 were heat treated.

TABLE 2 (*continued*)

Specimen			Stock		Coil		Section		Press. Sens.	Puls. Stress	Life	Parameters			
p	#	h	Mtl.	R_o	ø	a	b	$\frac{-\Delta ø}{ø}\frac{E}{p}$	$10^{-3}S$	$10^{-3}N$	$\frac{Rh}{a^2}$	$\frac{b}{a}$	$\frac{a^2}{10^2h^2}$	$\frac{10^6bh^3}{a^4}$
1000	C 231	46	CM	2.25*	300*	238 x	94 E	180	180	10	1.8	.39	.27	2810
1000	C 232	43,45	CM	"	"	240 x	93 E	190	197	10	1.7	.39	.30	2400
3000	C 233	24*	BC	.841*	255*	127*x	48*E	-	-	-	1.2	.38	.28	2400
3000	C 234	24*	BC	.837*	252*	125*x	50*E	-	-	-	1.3	.40	.27	3000
300	C 235	18*	PB	.848*	249*	188*x	52*-	-	-	-	.44	.28	1.09	260
200	C 236	16*	PB	.842*	249*	189*x	53*E	-	-	-	.36	.28	1.40	154
200	C 237	20*	Bz	.845*	249*	250*x	55*E	-	-	-	.26	.22	1.56	105
200	C 238	19*	Bz	1.13*	220*	183*x	49*E	-	-	-	.64	.27	.93	297
200	C 239	16*	Bz	.663*	219*	252*x	61*F	-	-	-	.17	.24	2.48	61
400	H 240	14*	Bz	.850*	600*	117*x	31*F	-	-	-	.84	.26	.70	400
100	C 241	10*	Bz	.742*	257*	187*x	27*E	-	-	-	.20	.14	3.5	19
65	C 242	8*	Bz	.800*	257*	185*x	47*E	-	-	-	.19	.25	5.4	20
25	C 243	8*	Bz	.829*	275*	298*x	59*-	-	-	-	.07	.20	13.9	3.1
20	C 244	10*	Bz	.829*	275*	297*x	58*-	-	-	-	.09	.20	7.6	6.3
275	H 245	6*	S	.390	See	99*x	27*F	2000			.24	.27		61
2000	H 246	15.7*	S	.390	Note	109*x	35*F	330			.52	.32		960
3000	H 247	16*	S	.390	Below	107*x	37*F	275			.54	.35		1155
4360	H 248	20.5*	S	.380		109*x	37*F	131			.66	.34		2260
9250	H 249	26.5*	S	.380		111*x	42*F	62			.82	.38		5150
100	H 250	12	PB	.781	1800	185 x	27 F	3440			.27	.15	2.38	40
200	H 251	16	PB	.781	1800	185 x	25 F	1890			.37	.14	1.34	88
400	H 252	20	PB	.781	1800	182 x	23 F	1030			.47	.13	.83	168
600	H 253	24	PB	.781	1800	180 x	20 F	665			.58	.11	.56	263
1000	H 254	25	BC	.781	1800	170 x	20 F	686			.68	.12	.46	374
300	H 255	16	CRS	.781	1800	180 x	15 F	2335			.39	.08	1.26	58
800	H 256	24	CRS	.781	1800	175 x	20 F	654			.61	.11	.53	295

231,232 are chrome molybdenum steel; assume $E=28.5\times10^6$ psi; $\Delta ø/ø$ based on Δy; inner wall .045, outer wall .043. # 54 to 232 were pulsed at 120 cpm, with arrangements to prevent tubes from heating at tip end. Internal radius at end of major diameter of section is changed by suppliers when forming tube, to give performance desired.

#245-249 from Wolf (9), 30 to 40 turn helices.

#250-253 are phosphor bronze grade A: Cu 95, Sn 5, Pb .035; $E=15.5\times10^6$; Rockwell B80. Formed and has silver soldered fittings; then stress relieved 8 hours at 300°F.

#254 is beryllium copper: Be 1.70 to 1.90, Ni+Co .60, balance Cu. $E=18.5\times10^6$; 1/4 hard (Rockwell B 70 to 75). Formed and has silver soldered fittings; then is heat treated 3 hours at 600°F. Hardness after heat treatment, Rockwell C 35.

#255,256 are type 316 stainless steel: Cr 17, Ni 13, Mo 2.5. $E=28\times10^6$; Rockwell B 80. Formed and has welded fittings; then is stress relieved 8 hours at 600°F.

Acknowledgment

For the painstaking work of reducing these data and plotting the graphs, acknowledgment is due in part to Mr. Jenny of the subcommittee, and in larger part to Leonard Mordfin, Robert Czorny, and Nazz Viola, of the Engineering Mechanics Section, National Bureau of Standards, whose assistance was secured by financial support to RCMPE from Engineering Foundation, Inc.

Thanks are due J. R. Keim, secretary of RCMPE, for the correction of major errors in Table 1 of the preprint.

Bibliography

1 "Die Bewegungslehre von Rohrenfedern," by W. Wuest, *Zeitschrift für Instrumentenkunde*, vol. 63, 1943, pp. 416–428.

2 "A Dimensional Analysis Approach to Bourdon Tube Design," by K. Goitein, *Instrument Practice*, vol. 6, 1952, pp. 748–755.

3 "Theories on Pressure Measurement With Bourdon Tubes," by F. B. Jennings and J. T. Duane, unpublished report, General Electric Company.

4 "Bibliography on Bourdon Tubes & Bourdon Tube Gages," by L. Van der Pyl, ASME Paper 53—IRD-1, September, 1953.

5 "Theories on Bourdon Tubes," by F. B. Jennings, report to RCMPE, June 21, 1954, and ASME paper 54—A-168, published in this issue, pp. 55–64.

6 "Selecting and Working Bourdon-Tube Materials," by J. B. Giacobbe and A. M. Bounds, *Instrument Manufacturing*, vol. 20, 1952, pp. 130–132, 136–137.

7 "Stress and Deformations of Toroidal Shells of Elliptical Cross Sections," by R. A. Clark, T. I. Gilroy, and E. Reissner, *Journal of Applied Mechanics*, Trans. ASME, vol. 74, 1952, pp. 37–48.

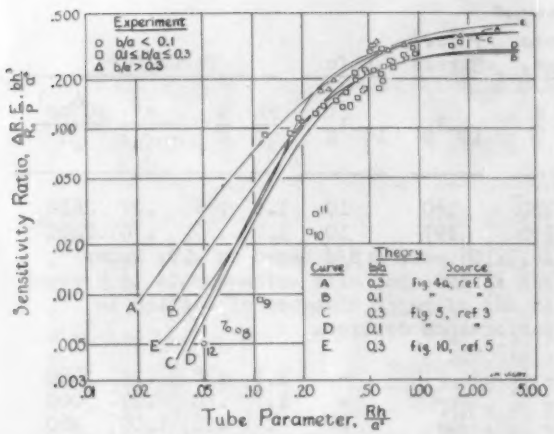


FIG. 6 PRESSURE SENSITIVITY OF FLAT OVALS AND DEES

8 "The Influence of the Shape of the Cross-Section on the Behavior of Bourdon Tubes," by W. Wuest, *Ingenieur-Archiv*, vol. 20, 1952, pp. 116–125, translated by R. P. Jerrard, General Electric Company, ASME Paper 54—A-165.

9 "An Elementary Theory of the Bourdon Gage," by Alfred Wolf, *Journal of Applied Mechanics*, Trans. ASME, vol. 68, 1946, p. A-207; discussion, *Journal of Applied Mechanics*, Trans. ASME, vol. 69, 1947, p. A-165.

Discussion

J. B. GIACOBBE.[3] The author is to be commended for bringing together in this excellent paper the theory and practice of Bourdon tubes which for too long a period of time have been separated by much too wide a gap. In fact, there are more than a few Bourdon gage manufacturers who will readily admit that there is considerably more art than science in the making of Bourdon gages. Happily, they are just as eager as the researcher to develop ground rules and formulas which will reduce trial and error to a minimum.

That the art should be so far ahead of the science is perhaps understandable when one considers the complexities of the Bourdon-tube theory. As pointed out by the author, the uncoiling of a pressure-gage element is proportional to coil angle ϕ, internal pressure p, reciprocal modulus of elasticity E^{-1}, and geometrical parameter denoted as k_1. That is a simple enough statement of fact, but there still remains to be determined just how this important parameter is related to tube geometry. In Figs. 3 and 6 of the paper, k_1, expressed as the sensitivity ratio

$$\frac{\Delta R}{R_o}\frac{E}{p}\frac{bh^3}{a_4}$$

is plotted against a tube parameter, Rh/a^2, and as will be recalled, the experimental data fall fairly close to the theoretical curves. This is very encouraging and indicates that the researchers have come quite close to defining the k_1. Indeed, when it is considered that the data plotted in Fig. 3 for elliptical cross sections and those in Fig. 6 for flat oval cross sections represent several different alloys both ferrous and nonferrous, heat-treated and nonheat-treated, as well as different manufacturers, it is surprising to find that the agreement is as good as it is.

However, there is little doubt that the scatter reflects in large measure the manufacturing variables alluded to. Each Bourdon-gage producer has his own specification for Bourdon tubing with respect to cross section, "temper" or hardness, composition and, in addition, has his own method of fabrication and heat-treatment, if any. It is to be hoped that some standardization, particularly with respect to cross section, can be achieved in time to come but the other variables mentioned probably account in large measure for the erratic nature of the data. An investigation or test program should be set up by the Theory Subcommittee eliminating as many of these variables as possible in order to check theory.

It may be argued that a sound theory of the Bourdon-tube element will not involve certain metallurgical factors relating to the type of metal or alloy being used. However, from a practical standpoint the engineer wants a formula which will work at least to a first approximation. Therefore it might be worth while to divide the metals and alloys used for Bourdon tubing into at least three groups which may hasten the development of a practical working formula. These groups are (*a*) copper-base alloys, such as phosphor bronze and beryllium copper, (*b*) nickel-base alloys, such as Ni Span C and K Monel and (*c*) iron base such as chrome-molybdenum steel and AISI Type 403 stainless steel. This will have the effect of separating the alloys into three different groups with respect to elastic moduli which property does have an important effect on pressure sensitivity. This, together with the foregoing recommendations, should produce some interesting results.

The author is requested to supply information on the following:

1 Should corner radius have any appreciable effect on pressure sensitivity?

2 Is the Wuest curve in Fig. 3 independent of axes ratio?

3 It is assumed that stress concentration exists at the corner radius of any cross section since most fatigue failures occur in this area. In measurement of stress, therefore, is account taken of the concentration of stress at the corners?

K. GOITEIN.[4] Everyone who in the past has tried to tackle this intriguing problem of Bourdon tubes will agree with the writer that the paper presented here marks a decisive step in the direction of obtaining useful results from the combined efforts of the scientific investigator and the practical engineer.

It is particularly gratifying that experimental results have been presented in a standardized form so that comparisons between different test results and between these and various theories become possible for the first time.

The large number of individual test results that have been tabulated presents a formidable problem of evaluation and it is regrettable that many of the specifications quote "nominal" dimensions only. An inaccuracy of 5 to 10 per cent, say, in the wall thickness might easily amount to sensitivity errors of 30 per cent or more and any comparison with theory becomes rather illusory under these circumstances.

There can be no doubt at all that much of the "scatter" in the graphs, Figs. 3 and 6, is due to inaccurate testing conditions, different methods of manufacture, like heat-treatment, end effects due to soldering and welding, differences in the estimation of Young's modulus, and other such imponderables. To obtain really valid comparisons with theory, or to establish a useful empirical approximation, all these conditions will have to be controlled carefully in a series of scientific tests.

But even allowing for all these explanations of scatter, it appears that theory and practice have not yet come together. Representation on log paper makes discrepancies appear small but

[3] Plant Metallurgist, Superior Tube Company, Norristown, Pa. Mem. ASME.

[4] Elliott Bros., Ltd., Century Works, Lewisham, London, S. E., England.

it cannot disguise the fact that a factor of 2 is quite frequent and some results differ from theory by a factor of as much as 5. Scatter, in some cases, goes up to a factor of 10.

The writer believes that this is more than can be explained by inaccurate testing conditions. The theories are admittedly approximative. They obviously require a great deal of improvement to be of practical value. This may take a lot of time and the instrument designer cannot wait until better results are available. There is, therefore, an immediate need to represent experimental results in a digestible and systematic form.

For this purpose the tube parameter and sensitivity ratio, as defined in theory, are rather clumsy instruments of presentation. The fact that we have to face is that it is not possible to represent the sensitivity of a Bourdon tube as a function of one parameter, or even of two parameters. In fact, for a given shape of cross section, say, flat oval or elliptic, three parameters, or dimensionless ratios are required to define the geometry. Any three ratios formed from the four dimensions R, a, b, and h will completely characterize a Bourdon tube, or rather a group of geometrically similar tubes, and a complete representation, not prejudiced by any approximative theories, will have to represent the pressure sensitivity as a function of three parameters thus

$$\frac{\Delta R}{R}\frac{E}{p} = f\left(\frac{R}{h}, \frac{b}{a}, \frac{R}{b}\right)$$

This is, of course, a four-dimensional curve not capable of direct graphical presentation. The nearest approach would be a series of curve families, or an equivalent tabulation scheme. It is of course possible, and indeed quite likely, that test results, tabulated in this form, will be found to be capable of very close approximation, within limited ranges, by simple power functions, as the writer has found in one series of tests (see reference 2 of author's bibliography). This would immensely simplify practical prediction of Bourdon-tube deflections.

Some minor points in the paper require clarification:

The author mentions a change of 1 to 2 per cent of bend radius. In many practical cases this can go up to 5 or 6 per cent, when the formula $\Delta R/R = -\Delta\varphi/\varphi$ becomes inaccurate by about this amount. Would it not be advisable to use the quantity $\Delta R/R_0$ which is independent of the actual amount of deflection, in place of $\Delta\varphi/\varphi$ which is not, in any sensitivity formula?

WALTER WUEST.[5] During the war the writer made a similar investigation covering a somewhat more extended range of data.[6] In the beginning of the investigation he also found a large scattering of the measured values with respect to theoretical data. It was found that the scatter was essentially caused by inappropriate methods of measuring. For this reason a special device was developed for measuring the sensitivity of thin-walled Bourdon tubes, containing a movable mirror, which did not exert any force on the tube.

In a later report,[7] a thorough investigation of flat spiral steel springs was undertaken. The partial sensitivity of each coil was measured and the values were found to lie within ±5 per cent of the theoretical ones. For the springs under investigation the sensitivity of the inner coil was in most cases only one third that of the outer coil. By measuring also the partial elastic hysteresis of each coil, the dependence of maximum stress on radius or tube parameter could be evaluated.

[5] Max-Planck-Institut für Strömungsforschung, Göttingen, Germany.

[6] Unpublished Report 97 of Schaffer & Budenberg, Ltd., February, 1943.

[7] Unpublished Report 145 of Schaffer & Budenberg, Ltd., December, 1943.

AUTHOR'S CLOSURE

The committee did not undertake separate plotting of material groups (*a*) and (*c*) as suggested by Mr. Giacobbe. The data of Table 2 are adequate to permit this, except that no data are available on nickel-base alloys. The questions Mr. Giacobbe raises can be answered as follows:

1 The effect of "corner radius" on pressure sensitivity is not provided for in the theories quoted, although the sixth-order polynomial of Wuest tends to do so. From the practical viewpoint, one user reports that his supplier is able to alter pressure sensitivity by changing corner radius.

2 The Wuest curve of Fig. 3 is independent of the semiaxes ratio b/a, provided b/a is "small."

3 Stress concentration at the corner radius is ignored in the theories quoted. The few measured values reported are based on Stress-coat readings on the external surface at the major axis.

Mr. Goitein is obviously correct in pointing out the need for careful control of as many variables as possible. Table 2, while not a series of scientific tests, does represent the first real cooperative effort by the industry, and is intended to blaze the trail for more precise experimentation.

For trial of the three geometric parameters R/h, b/a, R/b, the data of Table 2 are available, but are perhaps too imprecise to warrant the labor of a least-squares analysis. A couple of trials with the parameters and exponents of reference (2) indicate that those values recommended for beryllium copper are not suitable for trumpet brass.

As a nondimensional measure of deflection, some theories use $\Delta R/R$, others use $-\Delta\phi/\phi$, others assume their equivalence. There seems to be no theoretical reason why either ratio should vary with coil angle. The development of Equations [1], [2], and [3] appears to be correct in showing R and ϕ without the zero subscript, although of course there are other approximations involved. However, in the computations for Table 2 and Figs. 3 and 6, values of R_0 and ϕ_0 were the ones actually used, resulting in imprecisions on the order of $\Delta R/R$.

A related matter of some interest is the fact that most of the pressure-sensitivity values of Table 2 were computed from measured travels Δy or $(\Delta x^2 + \Delta y^2)^{1/2}$. Experimentally, tip travel is easier to measure precisely than is either ΔR or $\Delta\phi$, and is of more direct interest. However, for specimens 31 to 42, four distinct sets of measurements were taken, giving independent k_1 values as follows:

TABLE 3

Serial	Lab A $(\Delta x^2 + \Delta y^2)^{1/2}/R_0$	Lab B $(\Delta x^2 + \Delta y^2)^{1/2}/R_0$	Lab B $\Delta R/R_0$	Lab B $\Delta\phi/\phi_0$
31	1410	1410	1370	1360
32	1380	1405	1610	1110
33	1340	1450	1250	..
34	1240	1230	..	..
35	498	513	525	575
36	425	435	404	540
37	485	540	634	488
38	390	425	451	352
39	268	272	324	345
40	302	306	320	346
41	276	278	316	286
42	236	255	238	185

Our measurements of pressure sensitivity were made, as were Herr Wuest's later ones, by optical means which did not impose any force on the tube. We hope that his results can now be made available to RCMPE. In view of the extremely small values to be expected for total elastic hysteresis, often on the order of 0.1 or 0.2 per cent (and occasionally appearing to be negative), we should also be interested in his techniques for measuring partial elastic hysteresis.

Studies in Cold-Drawing

Part 3, Determination of Friction Coefficient[1]

By H. MAJORS, JR.,[2] UNIVERSITY, ALA.

Part 3 is a continuation of the studies in cold-drawing.[3] Average coefficients of drawing friction were determined for forward and reversed drawing of 2S-O aluminum and SAE 1020 steel by Sachs' theory and compared with measurements using SR-4 wire strain gages on the outer die surface to determine die pressures. Die profiles were determined by plastic castings. A series of dies were used to provide true reductions of rod area up to 0.76 at drawing speeds of 2 ipm.

Literature Review

THERE is an abundant literature on wire-drawing, but until recently, the predominant concern has been with the relationships of drawing force and power on the various conditions of a rod passing through a die.

One of the earliest articles is that of Smith (1)[4] in 1887. Except for Musiol's (2) work in 1900 no discussion appears on wire-drawing until that of Lewis (3, 4, 5) in 1915, and then later in 1933 and 1938, where he attempted to relate power used with the dimensions and properties of the wire by empirical factors. Except for Elder's (6) extension of Lewis' work to include strain hardening, little occurs in American literature up to 1944 regarding wire drawing, especially forces in dies. Following 1944, one finds several articles on the theory of wire-drawing in American and English literature.

In England, Alkins and Cartwright (7) investigated relative reductions of concentric copper tubular layers built on a solid core. Alkins (8) reports tests conducted on 0.435-in-diam annealed copper rod and found that when the reduction is effected by successive drafts, the tensile strengths attained for any given reduction of the original area are lower, the lighter the drafting, that is, the more numerous the stages by which the reduction is carried out. Dies with a straight taper of 5 deg were lubricated with vegetable oil when the short rods were pulled. Up to 15 per cent nominal reduction of area, the sequence had no effect on a plot of nominal tensile strength versus reduction of area; from a reduction of area at 15 per cent and beyond there was no longer a linear relationship and the tensile strengths were higher for single drafts than five successive drafts. It was lowest for nine successive drafts. No "reverse" sequence was investigated. Alkins concluded that there was a fundamental difference between reductions by cold-drawing of less than 15 per cent and those which exceed this amount.

Francis and Thompson (9) give an account of the drawing of nonferrous wires approximately 0.08 in. diam. Several conclusions were drawn from tests on aluminum, copper, bronze, 70:30 brass, 7 per cent nickel-brass, and 20 per cent nickel-brass as follows:

1 Approximately, the drawing force required to draw wire is proportional to the tensile strength of the original material and independent of the speed of drawing over the range from 20 to 600 fpm.

2 A plot of drawing force versus nominal reduction of area is linear. For perfect lubrication the curve passes through the origin, whereas in other cases the force is increased for all reductions by a constant amount.

3 For a 20 per cent reduction of area the minimum load is required when a plain conical die with semiangle of $5^1/_2$ to 6 deg is used. For a 35 per cent reduction there is very little difference from $5^1/_2$ to 12 deg, but with angles smaller than $5^1/_2$ deg the power needed rises very rapidly. In all cases a die with a radial taper required a considerably greater tension than does the best of the straight-taper dies.

4 Best lubrication was obtained with a sodium soap containing not less than 75 per cent saponified fatty acids, and as free as possible from water and glycerol. Surface finish was better with liquid than with solid lubricants.

5 Using tungsten-carbide dies the coefficient of friction is reduced with soap certainly to less than 0.04, which is less than a quarter of that previously determined for steel dies when drawing mild steel.

6 With diamond dies and liquid lubricants less power was required than when carbide dies were used.

Lewis in a discussion of the foregoing paper comments that the coefficient of friction is irregular in the first draft and higher while the film is being established than in subsequent drafts.

Many attempts have been made to calculate the force required to draw a wire through a die. One of the early attempts at a mathematical analysis was that due to Sachs (10). Korber and Eichinger (11) improved on Sachs' work by considering the direction of the greatest principal strains as directed toward the die-cone vertex, as well as Davis and Dokos (12) who considered strain-hardening effects. Lunt and MacLellan (13) and MacLellan (14) have discussed and generalized the theory originated by Sachs.

An attempt was made by Godfrey (15) to show the effect of a change in operation upon the physical properties of 0.67 per cent patented wire-drawn 40, 30, 20, and 10 per cent reduction per draft; also upon a 0.27 per cent carbon patented rod drawn 40, 30, 20 per cent reduction per draft. The 0.67 per cent carbon was affected to a greater extent by the rate of drawing than the 0.27 per cent carbon wire; thus, the greater the reduction in area per draft, the higher is the nominal tensile strength for the same total amount of drawing. The tensile strength of cold-

[1] Funds for this investigation were made available by the Research Corporation, New York, N. Y., during the period of December, 1949, to February, 1952, while the author was on the staff of the University of Alabama. F. J. Mehringer examined the manuscript and offered many suggestions.

[2] Research Professor of Engineering Mechanics, Bureau of Engineering Research, University of Alabama. Mem. ASME.

[3] "Studies in Cold-Drawing—Part 1, Effect of Cold-Drawing on Steel," by H. Majors, Jr., Trans. ASME, vol. 77, 1955, pp. 37–47. "Part 2, Cold-Working 2S-O Aluminum," ibid., pp. 49–56.

[4] Numbers in parentheses refer to the Bibliography at the end of the paper.

Contributed by the Metal Engineering Division and presented at the Annual Meeting, New York, N. Y., November 28–December 3, 1954, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, April 29, 1953. Paper No. 54—A-114.

drawn wire was reduced by the machine-straightening operation and its ductility was increased as measured by the elongation.

Experimental Work

The objectives of the experimental work were as follows:

1 To cold-draw 1-in-diam rods of SAE 1020 steel and 2S-O aluminum through dies in various sequences or passes in order to note the variation of the coefficient of friction.

2 To investigate the use of SR-4 strain gages on the outer periphery of the dies in order to measure the average die pressure on the rod.

Conditions of the tests are reported in Parts 1 and 2.[3] Table 1 defines the sequences used while Tables 2 and 3 give the physical dimension of the dies used to cold-draw the rods. Fig. 1 indicates the die shape. Reproductions of the die surface and shape were made by using a polyester resin as described by Sawyer (16),[3] Fig 2.

Sachs (21) computed friction coefficients from a knowledge of the ideal drawing stress, actual drawing stress, value of yield strength and the die dimensions. His basic equation was one of statics, summing forces along the rod axis. Die angles of 2, 4, 8, 16, and 32 deg were investigated. All dies had an average value of about 0.08 for the friction coefficient.

[3] The plastic is packaged as Marcokit No. 249 (17) which was developed originally to provide an improved and simplified method for checking machined surfaces. Manufactured by Marco Chemicals, Inc., Sewaren, N. J.

A direct determination of the coefficient of friction under drawing conditions can be made if the mean pressure on the die is known. MacLellan (18) attempted this from experiments with "split" dies and indicated such results would be independent of plasticity theory. Tests were conducted on wire at drawing speeds of 12 ipm using a 6-deg tungsten-carbide die. The results were not satisfactory which was believed to be due to the penetration of the lubricant between the halves of the die.

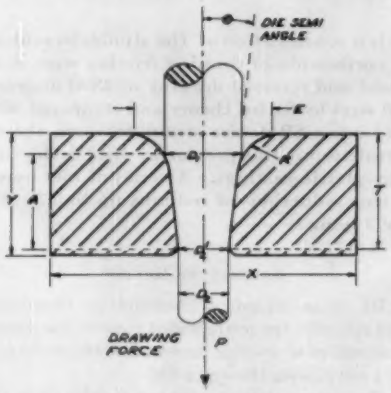


Fig. 1 Nomenclature Used in Drawing Rod—See Tables 2 and 3

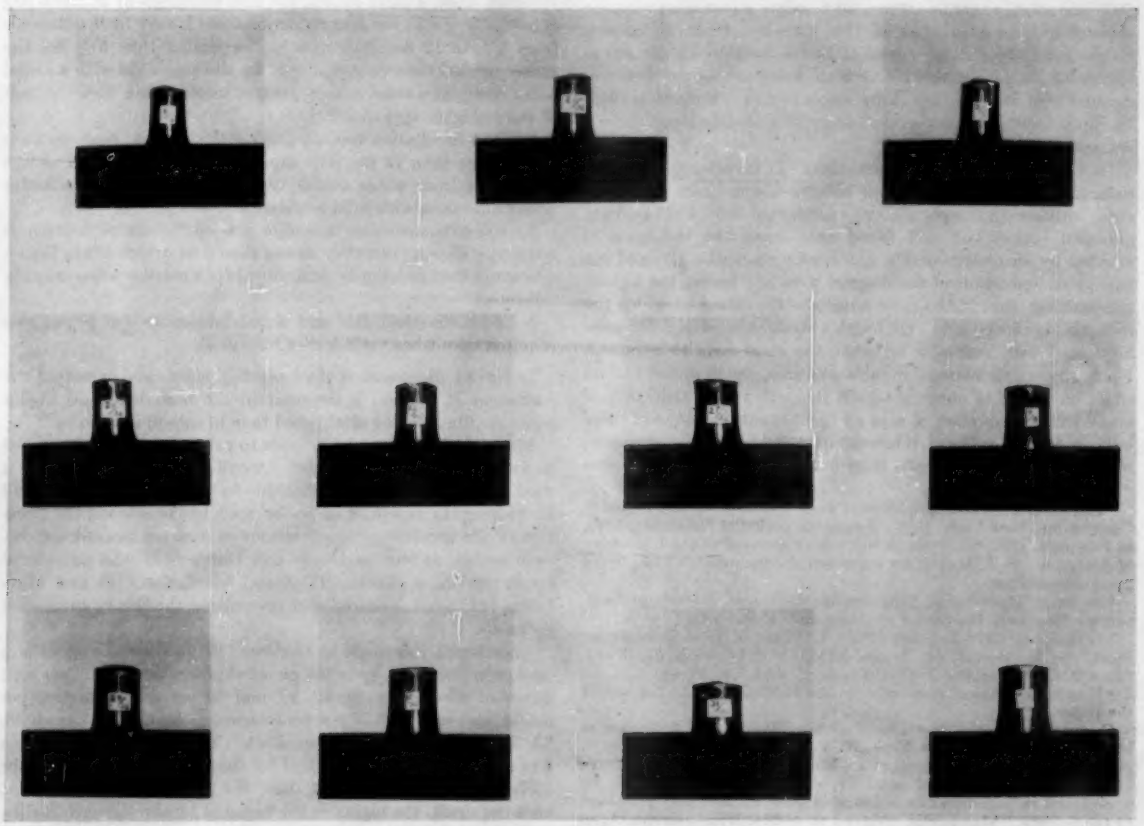


Fig. 2 Photograph of Plastic Die Castings

Because large rods were used in this investigation, it was considered feasible to use SR-4 bonded wire resistance strain gages placed tangentially on the outer surface of the dies. Furthermore, in large rods the more strain-hardened surface layers are restricted to a smaller proportion of the cross section than in thinner wire. The large drawing force should result in a large pressure of the rod on the die. In contrast to the split-die method (18) of measuring the die force, this method measures the average pressure. Equation [1] relates drawing stress to the mean die pressure (see Appendix)

$$\left[\frac{\frac{\sigma_x}{p}}{\left(\frac{A_1}{A_2}-1\right)}\right]-1=\mu\cot\theta \qquad [1]$$

where

σ_x = drawing stress on final area, psi
p = mean die pressure, psi
A_1 = initial rod area, sq in.
A_2 = final rod area, sq in.
θ = half die angle
μ = mean coefficient of friction

TABLE 1 EXPLANATION OF SEQUENCES IN DRAWING RODS

NOMINAL DIE SIZE INCHES	1	63/64	31/32	30/32	29/32	28/32	27/32	26/32	25/32	24/32	23/32	22/32	
NOT REVERSED NOT BY-PASSED	A	A	A	A	A	A	A	A	A	A	A	A	A
REVERSED NOT BY-PASSED	A	B	A	B	A	B	A	B	A	B	A	B	B A
NOT REVERSED BY-PASSED	A			A		A		A		A		A	A
REVERSED BY-PASSED	A			B		A		B		A		B	B A

LETTERS INDICATE THE ROD END ENTERING THE DIE
BLANK SPACES INDICATE THE ROD DID NOT PASS THROUGH THAT DIE

TABLE 2 DIMENSIONS OF DIE

Nominal Size Of Die	D_2'	T	E	R	X	Y
1"	0.997	1.2	1.27	2.20	2.9	1.3
63/64	0.970	1.0	1.20	2.37	2.9	1.4
31/32	0.956	0.87	1.10	1.20	2.8	1.1
30/32	0.921	1.12	1.18	1.90	2.9	1.4
29/32	0.887	1.09	1.14	1.08	2.9	1.4
28/32	0.869	1.07	1.03	2.13	2.9	1.3
27/32	0.833	1.10	1.13	1.20	2.8	1.3
26/32	0.799	0.96	1.01	1.09	2.9	1.3
25/32	0.766	0.98	1.09	0.50	2.8	1.3
24/32	0.738	1.06	1.00	0.68	2.9	1.3
23/32	0.713	1.14	0.95	0.83	2.9	1.3
22/32	0.675	0.98	0.99	0.52	2.9	1.3

The dies with wire strain gages were calibrated with oil pressure supplied by a dead-weight tester according to Fig. 3, in which the oil enters the annular space between the inner die surface and a steel plug through the high-pressure fitting in the plug. O-rings prevented the escape of oil and held the oil pres-

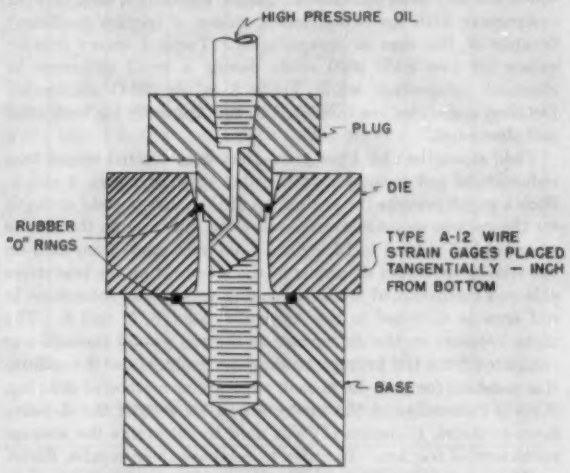


FIG. 3 METHOD FOR APPLYING OIL PRESSURE TO INNER DIE SURFACE IN ORDER TO CALIBRATE WIRE STRAIN GAGES PLACED ON OUTER SURFACE, TANGENTIALLY

TABLE 3 DIMENSIONS A, D_1, D_2, AND θ, FOR THE VARIOUS SEQUENCES

Nominal Die Size	SEQUENCE IN DRAWING											
	Not Reversed Not By-Passed				Reversed Not By-Passed				Not Reversed By-Passed				Reversed By-Passed			
	A	D_1	D_2'	TAN	A	D_1	D_2'	TAN	A	D_1	D_2'	TAN	A	D_1	D_2'	TAN
1"	.76	1.000	.997	.0020	.76	1.000	.997	.0020	.76	1.000	.996	.0020				
63/64"	.82	.997	.970	.0164												
31/32"	.72	.978	.956	.0153	.80	.997	.956	.0256								
30/32"	.90	.966	.921	.0250					1.00	.997	.921	.0300				
29/32"	.97	.930	.887	.0221	1.03	.966	.887	.0384								
28/32"	.90	.898	.869	.0161					1.00	.930	.869	.0305				
27/32"	.82	.872	.833	.0238	.90	.898	.833	.0361								
26/32"	.88	.842	.799	.0284					.95	.872	.799	.0384				
25/32"	.80	.808	.766	.0263	.95	.842	.766	.0400								
24/32"	.68	.775	.738	.0273					.85	.808	.738	.0413				
23/32"	.76	.747	.713	.0223	.95	.775	.713	.0327								
22/32"	.64	.719	.675	.0343	.84	.719	.675	.0343	.78	.747	.675	.0462				

sure steady for reading the strain-gage indicator. Each die was calibrated five times and the pressure in lb per sq in. was plotted against strain-gage differences in microinches. Pressures were applied to 10,000 psi. After determining the calibration constant for the die by applying oil pressure over the same length of die surface as was touched by the rod, various rods were pulled in different sequences and the corresponding indicated strains on the outside die surface were noted every 30 sec during the drawing operation. Average pulling loads and indicated strains were computed for all operations. Once the drawing operation reached a steady rate, drawing loads and pressures varied only slightly.

Discussion of Results

A summary of the kinetic coefficients of friction appears in Tables 4 and 5 where a comparison is made with those values obtained from pressure measurements, Equation [1], with those computed from Sachs' equation. Table 16 in the Appendix compares Sachs' equation with that of Davis and Dokos and shows the very close agreement. Sachs' equation is used here for comparison with the experimental values of friction coefficient because of the ease in computation. Table 4 shows friction values for two SAE 1020 steels having a small difference in chemical compositon while Table 5 is for 2S-O aluminum. Detailed quantities are tabulated in the Appendix for both steel and aluminum.

Yield strengths at 0.1 per cent offset were plotted versus true reduction of rod area as shown in the Appendix. Figs. 4 and 5. Such a graph permits the evaluation of the average yield strength for the various sequences or passes of the rod through the dies as recorded in Tables 6 to 12. Also, plots were made from data on the true stress-strain diagrams of the intercept on the true stress axis and coefficient of strain hardening versus true reductions in rod area as sketched in the Appendix, Figs. 6, 7, and 8. The mean pressure on the die surface as the rod passed through was computed from the average strain-gage readings and the calibration constant for each die in each sequence or method of drawing. With a knowledge of the pressure and converting the drawing force to stress, Equation [1] was used to determine the average coefficient of friction. In order to compare these results, Sachs' well-known equation on rod-drawing (10) was used to compute the coefficient of friction using average values of the yield strength at 0.1 per cent offset before and after the die. Five computation trials at least were needed for each determination.

One observes in Tables 4 and 5 that the friction values obtained from pressure measurements were higher than those obtained from Sachs' equation, except in a few instances for aluminum. A partial explanation of this was found from an examination of the die profiles, noting where the rod enters the die. Those rods that were cold-worked a short distance at the die entrance due to radial taper and then cold-worked with a straight tapered surface had higher coefficients of friction than when cold-worked entirely with a straight tapered die. For those rods that passed through every die in this study, it was noted that the $^{31}/_{32}$, $^{30}/_{32}$, $^{29}/_{32}$, $^{28}/_{32}$, and $^{24}/_{32}$-in. dies provided conical surfaces throughout the passage, whereas the other dies had a bell-shaped entrance for a short distance, then a conical surface for the completion of the draw; whereas the by-passed sequences showed higher coefficients from pressure measurements than from Sachs' equation in all cases except an instance or two for aluminum, since the cold-working was partially in the bell-shaped area and completed on the conical surface. Francis and Thompson (9) report that bell-shaped dies require considerably more tension than straight tapered dies. A larger drawing force is reflected in the equations by a larger coefficient of friction.

Another possible explanation, although it is a small influence,

TABLE 4 COMPARISON OF EXPERIMENTAL METHOD WITH THE USE OF SACHS' EQUATION TO DETERMINE THE COEFFICIENT OF FRICTION

NOMINAL DIE SIZE (IN.)	COEFFICIENT OF FRICTION							
	S.A.E. 1020 STEEL ROD No. 1				S.A.E. 1020 STEEL ROD No. 2			
	NOT REVERSED NOT BY-PASSED		REVERSED NOT BY-PASSED		NOT REVERSED BY-PASSED		REVERSED BY-PASSED	
	EXP.	SACH'S	EXP.	SACH'S	EXP.	SACH'S	EXP.	SACH'S
1	.08	.02	.08	.02	.08	.02	.08	.02
63/64	.12	.11						
31/32	.15	.15			.11	.07	.09	.07
30/32	.11	.10						
29/32	.09	.10			.15	.09	.12	.08
28/32	.14	.09						
27/32	.17	.14			.12	.07	.13	.08
26/32	.12	.10						
25/32	.22	.12	.14	.07	.16	.07	.16	.07
24/32	.15	.16	.13	.10			.11	.07
23/32	.15	.12	.17	.10	.12	.07	.17	.08
22/32	.20	.14			.17	.08	.16	.10

TABLE 5 COMPARISON OF EXPERIMENTAL METHOD WITH THE USE OF SACHS' EQUATION TO DETERMINE THE COEFFICIENT OF FRICTION

Nominal Die Size	COEFFICIENT OF FRICTION					
	2 S-O ALUMINUM					
	NOT REVERSED NOT BY-PASSED		REVERSED NOT BY-PASSED		NOT REVERSED BY-PASSED	
	EXP.	SACH'S	EXP.	SACH'S	EXP.	SACH'S
1	.05	.06	.10	.06	.09	.07
63/64	.13	.11	.11	.11		
31/32	.11	.10	.12	.10		
30/32	.08	.08	.08	.06	.05	.10
29/32	.09	.07	.08	.06		
28/32	.13	.07	.11	.06	.11	.07
27/32	.14	.10	.13	.09		
26/32	.10	.08	.15	.08	.15	.12
25/32	.16	.08	.17	.08		
24/32	.08	.07	.09	.08	.08	.10
23/32	.17	.08	.19	.09		
22/32	.13	.09	.13	.08	.09	.08

is that the deformation is heterogeneous, that is, nonuniform to some degree even in the large rods, causing an increase in cold-working more than that considered by Sachs' equation. The surfaces in contact with the die are under a large shear distortion. The direct measurement of the friction coefficient certainly reflects the effect of several real events including conditions of lubrication that were not optimum. As long as no metal adhered to the die for a particular method of lubrication, Parts 1 and 2,[3] no further change was made in lubrication procedures.

An examination of Tables 6 to 12 in the Appendix reveals that the pressures are high. One would suspect some difference

between the phenomena of friction at high pressures and at low pressure. Merchant (19) in a discussion to a paper by Dokos (20) computes values of the kinetic coefficient of friction which averaged about 0.23 for pressures of 20,000 to 50,000 psi under dry conditions when a mild-steel cylinder slides on an SAE 1045 disk at about 0.7 ipm. For sliding velocities from 0.7 to 600 ipm there was hardly any change in the coefficient. The conditions reported in this paper are believed to be for semilubricated conditions, hence one would expect lower friction values. Certainly, the Calgon coating and the cold-working filled many irregu-

TABLE 6 NOT REVERSED, NOT BY-PASSED SEQUENCE, SAE 1020 STEEL—ROD NO. 1

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure psi	Direct Determination of Coef. of Friction μ Equation (1)	Friction Coeff. μ Sach's Equation Using Average Yield Strength
1"	1.000	0.996	0.007	2,810	40,000	8,240	0.08	0.02
63/64	0.996	0.978	0.044	10,100	51,500	32,400	0.12	0.11
31/32	0.978	0.966	0.069	10,420	66,200	35,600	0.16	0.15
30/32	0.966	0.930	0.145	20,600	74,900	47,800	0.11	0.10
29/32	0.930	0.898	0.212	24,700	83,750	45,800	0.09	0.10
28/32	0.898	0.872	0.272	25,900	88,250	42,800	0.14	0.09
27/32	0.872	0.842	0.344	29,900	90,600	50,600	0.17	0.14
26/32	0.842	0.808	0.427	26,400	92,350	55,200	0.12	0.10
25/32	0.808	0.775	0.510	31,700	93,500	39,200	0.22	0.12
24/32	0.775	0.747	0.584	34,100	94,750	66,100	0.15	0.16
23/32	0.747	0.719	0.658	35,100	96,000	60,000	0.15	0.12
22/32	0.719	0.684	0.760	37,100	97,500	52,300	0.20	0.14

TABLE 7 REVERSED, NOT BY-PASSED SEQUENCE, SAE 1020 STEEL—ROD NO. 1

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00 in.	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure Psi	Direct Determination of Coef. of Friction μ Equation (1)	Friction Coeff. μ Sach's Equation using Average Yield Strength
1"	1.000	0.996	0.007	2,880	30,500	8,240	0.08	0.02
63/64	0.996	0.978	0.044	11,300	33,500			
31/32	0.978	0.966	0.069	11,630	37,000			
30/32	0.966	0.930	0.145	21,900	71,500			
29/32	0.930	0.898	0.212	21,200	76,500			
28/32	0.898	0.872	0.272	21,500	78,000			
27/32	0.872	0.842	0.344	23,000	79,000			
26/32	0.842	0.808	0.427	27,700	80,000			
25/32	0.808	0.775	0.510	21,200	80,200	39,000	.14	.07
24/32	0.775	0.747	0.584	23,200	81,500	53,900	.13	.10
23/32	0.747	0.719	0.658	30,700	84,000	45,600	.17	.10
22/32	0.719	0.684	0.760	31,900	90,500			

TABLE 8 NOT REVERSED, BY-PASSED SEQUENCE, SAE 1020 STEEL—ROD NO. 2

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00 in.	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure psi	Direct Determination of Coef. of Friction μ (1)	Friction Coeff. μ Sach's Equation Using Average Yield Strength
1"	1.000	0.996	0.007	2,820	40,000	8,240	0.08	0.02
63/64	0.996	0.978	0.004					
31/32	0.978	0.966	0.069	11,700	54,250	36,800	0.11	0.07
30/32	0.966	0.930	0.145					
29/32	0.930	0.898	0.212	31,900	77,500	40,700	0.15	0.09
28/32	0.898	0.872	0.272					
27/32	0.872	0.842	0.344	30,300	90,050	51,000	0.12	0.07
26/32	0.842	0.808	0.427					
25/32	0.808	0.775	0.510	38,500	96,500	42,000	0.16	0.07
24/32	0.775	0.747	0.584					
23/32	0.747	0.719	0.658	40,000	100,800	52,000	0.12	0.07
22/32	0.719	0.684	0.760	29,500	102,500	45,700	0.17	0.08

TABLE 9 REVERSED, BY-PASSED SEQUENCE, SAE 1020 STEEL—ROD NO. 2

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00 in.	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure psi	Direct Determination of Coef. of Friction μ Equation (1)	Friction Coeff. μ Sach's Equation Using Average Yield Strength
1"	1.000	0.996	0.007	2,880	40,000	8,040	0.08	0.02
63/64	0.996	0.978	0.044					
31/32	0.978	0.966	0.069	11,720	54,250	39,900	0.09	0.07
30/32	0.966	0.930	0.145					
29/32	0.930	0.896	0.212	30,700	76,350	48,000	0.12	0.08
28/32	0.898	0.872	0.262					
27/32	0.872	0.842	0.344	31,800	87,500	50,300	0.13	0.08
26/32	0.842	0.808	0.427					
25/32	0.808	0.775	0.510	35,900	94,000	40,300	0.16	0.07
24/32	0.775	0.747	0.584	24,000	96,000	60,200	0.11	0.07
23/32	0.747	0.719	0.658	31,600	99,800	48,400	0.17	0.08
22/32	0.719	0.684	0.760	33,400	101,000	56,100	0.16	0.10

arities, especially after the first cold-working pass. Also, there must be some metal-to-metal contact however small. It has been found in x-ray and electron-diffraction investigations that a type of unstable compound may be formed between metal and lubricant at the interface, in such a way as to function as a sliding cushion. In addition, it has been found that the power consumption for cold-work is not appreciably affected by the rate of deformation which would be so if there were complete liquid lubrication.

By comparing the values in Tables 4 and 5 with those obtained by other investigators, one does not find a serious disagreement in view of the varied conditions for lubrication. Sachs (21) presents values of the kinetic-friction coefficient under several circumstances. For steel rod on tungsten-carbide dies employing divers lubricants, values varied from 0.03 to 0.06. With brass on tungsten-carbide dies using rape oil, a value of 0.06 is reported; whereas, brass on soft-steel dies using rape oil, a value of 0.10 is mentioned. Williams (22) estimated the values to be between 0.07 and 0.29 for drawing copper and stainless steel on tungsten-carbide dies, with the lower value for wax and the higher for mineral-oil lubricant.

The influence of the manner of drawing (that is, whether forward or reversed drawing is applied and how many passes are needed to accomplish a given rod reduction) upon the coefficient of friction can be compared best by averaging the results in Tables 4 and 5, since an inspection reveals the values show no trends. On this basis there was no effect on the coefficient of friction due to straight or reverse drawing when steel rod was drawn. Using half the number of passes to accomplish a certain reduction (by-passed sequence) showed a slight reduction in the coefficient. The friction values for aluminum were slightly less than those for steel. Lower values were obtained on by-passed sequence than for the not-by-passed sequence. Direction of drawing had no effect.

TABLE 10 NOT REVERSED, NOT BY-PASSED SEQUENCE, 2S-O ALUMINUM

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00 in.	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure psi	Direct Determination of Coef. of μ Equation (1)	Friction Coeff. μ Sach's Equation Using Average Yield Strength
1"	1.000	0.996	0.007	885	3,750	4,010	0.05	0.06
63/64	0.996	0.978	0.044	1,625	5,880	4,850	0.13	0.11
31/32	0.978	0.966	0.069	1,400	9,080	6,830	0.11	0.10
30/32	0.966	0.930	0.145	2,860	10,350	8,520	0.08	0.08
29/32	0.930	0.898	0.212	3,510	12,030	6,880	0.09	0.07
28/32	0.898	0.872	0.272	3,600	13,200	6,580	0.13	0.07
27/32	0.872	0.842	0.344	4,280	13,580	8,850	0.14	0.10
26/32	0.842	0.808	0.427	4,160	14,350	10,130	0.10	0.08
25/32	0.808	0.775	0.510	4,360	15,050	7,250	0.16	0.08
24/32	0.775	0.747	0.584	3,710	15,580	11,750	0.08	0.07
23/32	0.747	0.719	0.658	5,220	16,000	7,740	0.17	0.08
22/32	0.719	0.684	0.760	5,040	16,500	10,000	0.13	0.09

TABLE 11 REVERSED, NOT BY-PASSED SEQUENCE, 2S-O ALUMINUM

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00 in.	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure psi	Direct Determination of Coef. of Friction μ Equation (1)	Friction Coeff. μ Sach's Equation Using Average Yield Strength
1"	1.000	0.996	0.007	873	3,750	1,870	0.10	0.06
63/64	0.996	0.978	0.044	1,520	5,880	5,150	0.11	0.11
31/32	0.078	0.966	0.069	1,415	9,580	6,160	0.12	0.10
30/32	0.966	0.930	0.145	2,580	11,000	8,020	0.08	0.06
29/32	0.930	0.898	0.212	3,100	12,300	6,770	0.08	0.06
28/32	0.898	0.872	0.272	3,180	13,080	6,630	0.11	0.06
27/32	0.872	0.842	0.344	3,990	13,750	8,430	0.13	0.09
26/32	0.842	0.808	0.427	4,290	14,300	6,830	0.15	0.08
25/32	0.808	0.775	0.510	4,510	14,900	6,870	0.17	0.08
24/32	0.775	0.747	0.584	4,130	15,480	12,540	0.09	0.08
23/32	0.747	0.719	0.658	5,200	15,900	6,940	0.19	0.09
22/32	0.719	0.684	0.760	4,780	16,300	9,820	0.13	0.08

TABLE 12 NOT REVERSED, BY-PASSED, 2S-O ALUMINUM

Nominal Die Size Inches	Diameter Before D_1 Inches	Diameter After D_2 Inches	True Reduction of Rod Area D_1=1.00 in.	Average Drawing Stress psi	Average Yield Strength psi	Average Pressure psi	Direct Determination of Coef. of Friction μ Equation (1)	Friction Coeff. μ Sach's Equation Using Average Yield Strength
1"	1.000	0.996	0.007	898	3,750	2,140	0.09	0.07
63/64	0.996	0.978	0.044					
31/32	0.978	0.966	0.069					
30/32	0.966	0.930	0.145	3,330	7,880	10,170	0.05	0.10
29/32	0.930	0.898	0.212					
28/32	0.898	0.872	0.272	4,730	12,900	7,650	0.11	0.07
27/32	0.872	0.842	0.344					
26/32	0.842	0.808	0.427	7,220	14,500	8,900	0.15	0.12
25/32	0.808	0.775	0.510					
24/32	0.775	0.747	0.584	6,750	15,500	13,490	0.08	0.10
23/32	0.747	0.719	0.658					
22/32	0.719	0.684	0.760	6,870	16,400	11,920	0.09	0.08

Conclusion

Parts 1, 2,[2] and 3 provide a full record of the mechanical properties of the steel and aluminum rods as well as the die profiles, measurements of drawing stress, and average die pressure. The effects of the manner of drawing or sequence were determined by using a series of hardened-steel dies in contrast to a series of rods. These results are for dies with half-angles of about 1 deg, in which

the average rod speed was about 2 ipm. By presenting a series of tension and torsion curves, hardness values, residual-stress measurements, some photomicrographs, impact and fatigue data, one has ample data to examine the several wire-drawing theories. Figs. 4 to 8 indicate the influence of the manner of drawing upon the true stress-strain properties.

The use of SR-4 wire strain gages on the outside die surface provides a method for recording the sum total of effects contributing to die pressure. Here is a tool to determine the effects of different lubricants and coatings upon the drawing force and die pressure. Dynamic recordings can be made by feeding the strain-gage output to a recording device such as the Brush or Sanborn equipment.

When Sachs' equation is used for computing the coefficient of friction, the sequence effect or manner of drawing had no appreciable effect on the coefficient. This was observed for the two steels as well as 2S-O aluminum. Values of the coefficient were similar whether large or small reductions of the rod were made. Frictional values were slightly less for aluminum than those for steel. The average of all dies for steel rod No. 1 was 0.10; for steel rod No. 2 it was 0.07; whereas for aluminum it was 0.08.

Frictional values determined from the pressure measurements using wire gages were consistently higher than those computed using Sachs' equation. Values of friction coefficient were lower for aluminum than for steel. Average values for steel No. 1 were 0.14; for steel No. 2 they were 0.13; whereas for aluminum they were 0.11.

Those dies that provided a conical surface throughout the cold-drawing process had friction coefficients that agreed well with Sachs' equation. Using Sachs' equation, the average for steel rods Nos. 1 and 2 was 0.10 and for 2S-O aluminum it was 0.08. Using pressure measurements, the average value of the friction coefficient for steel rod No. 1 was 0.11; for steel No. 2 it was 0.13; and for 2S-O aluminum it was 0.10.

TABLE 13 SUMMARY OF DRAWING LOADS REQUIRED TO COLD-DRAW SAE 1020 STEEL CYLINDRICAL RODS—ROD NO. 1
(Calgon coating—lubricant, paste of SAE 40 machine oil and Molycote powder.)

True Reduction of Rod Area	Nominal Die Size Inches	Average Drawing Load In Pounds	
		Not Reversed Not By-Passed	Reversed Not By-Passed
0.087	1	2,150	2,150
0.044	63/64	8,100	7,920
0.070	31/32	9,010	7,630
0.145	15/16	13,760	12,420
0.212	29/32	14,145	13,510
0.272	7/8	15,570	10,680
0.344	27/32	14,750	11,770
0.427	13/16	15,100	13,820
0.510	25/32	13,360	9,830
0.584	3/4	13,390	20,330
0.658	23/32	14,160	12,320
0.760	11/16	12,680	11,920

TABLE 14 SUMMARY OF DRAWING LOADS REQUIRED TO COLD-DRAW SAE 1020 STEEL CYLINDRICAL RODS—ROD NO. 1
(Calgon coating—lubricant, paste of SAE 40 motor oil and Molycote powder.)

True Reduction of Rod Area	Nominal Die Size Inches	Not Reversed, By Passed Sequence Average Drawing Load in Pounds			
0.007	1			2,240	2,240
0.044	63/64				
0.070	31/32	9,980			
0.145	15/16		18,700	15,500	22,050
0.212	29/32				22,050
0.272	7/8				
0.344	27/32			20,900	
0.427	13/16				20,300
0.510	25/32				
0.584	3/4				16,120
0.658	23/32				
0.760	11/16				

BIBLIOGRAPHY

1 "Flow of Metals in the Drawing Process," by O. Smith, *Journal of The Franklin Institute*, vol. 122, no. 5, 1886, pp. 321–346; vol. 123, 1887, p. 232.

2 "The Drawing in Extrusion Presses in Theory and Practice," by K. Musiol, *Kingler's Polytech. Journal*, vol. 315, 1900, pp. 428–432, 442–447.

3 "Power Requirements for Wire-Drawing," by K. B. Lewis, *Blast Furnace and Steel Plant*, vol. 3, 1915, pp. 1031–1034.

4 "The Wire-Drawing Die—Part 1," by K. B. Lewis, *Wire and Wire Products*, vol. 8, 1933, pp. 197–200; Part 2, pp. 234, 239, 243–251; Part 3, pp. 266–269; Part 4, pp. 331–333.

5 "Notes on Die Stresses," by K. B. Lewis, *Wire and Wire Products*, vol. 13, 1938, pp. 443, 476–477.

6 "The Wire-Drawing Die," by F. C. Elder, *Wire and Wire Products*, vol. 19, 1944, pp. 25–29, 32, 33.

7 "Experiments in Wire-Drawing. Part I—Behavior of a Composite Rod," by W. E. Alkins and W. Cartwright, *Journal of the Institute of Metals*, vol. 16, 1931, pp. 293–312.

8 "Experiments in Wire-Drawing, Part II," by W. E. Alkins, *Journal of the Institute of Metals*, vol. 46, 1931, pp. 304–312.

9 "The Drawing of Non-Ferrous Wires," by E. L. Francis and F. C. Thompson, *Journal of the Institute of Metals*, vol. 46, 1931, pp. 313–351.

10 "Die Plastizitat von Metallen," by George Sachs, *Zeitschrift*

TABLE 15 SUMMARY OF DRAWING LOADS REQUIRED TO COLD-DRAW SAE 1020 STEEL CYLINDRICAL RODS—ROD NO. 2
(Calgon coating—lubricant, paste of SAE 40 machine oil and Molycote powder.)

True Reduction of Rod Area	Nominal Die Size Inches	Average Drawing Load In Pounds			
		Not Reversed Not By-Passed	Reversed Not By-Passed	Not Reversed By-Passed	Reversed By-Passed
0.007	1	2,240	2,240	2,240	2,240
0.044	63/64	8,470	6,570		
0.070	31/32	8,840	8,520	8,570	8,580
0.145	15/16	13,850	14,890		
0.212	29/32	15,610	13,410	20,190	19,400
0.272	7/8	15,020	12,830		
0.344	27/32	16/680	12,810	16,850	17,710
0.427	13/16	14,580	14,230		
0.510	25/32	14,940	13,210	18,200	16,950
0.584	3/4	14,945	13,630		10,520
0.658	23/32	14,290	12,970	16,270	12,380
0.760	11/16	13,650	11,710	10,860	12,280

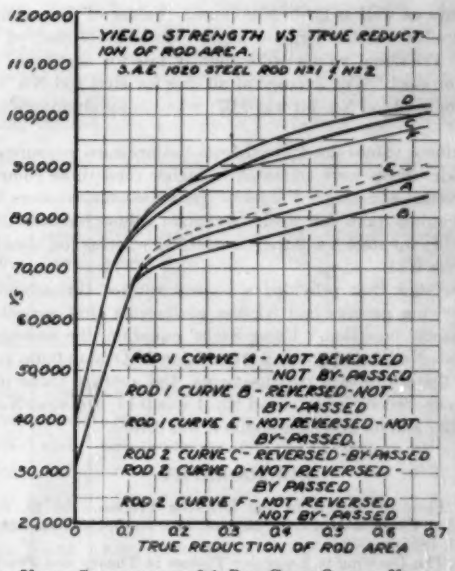

YIELD STRENGTH VS TRUE REDUCTION OF ROD AREA.
S.A.E. 1020 STEEL ROD NO 1 & NO 2
ROD 1 CURVE A - NOT REVERSED NOT BY-PASSED
ROD 1 CURVE B - REVERSED-NOT BY-PASSED
ROD 1 CURVE C - NOT REVERSED-NOT BY-PASSED.
ROD 2 CURVE C - REVERSED-BY-PASSED
ROD 2 CURVE D - NOT REVERSED - BY PASSED
ROD 2 CURVE F - NOT REVERSED NOT BY-PASSED
TRUE REDUCTION OF ROD AREA

FIG. 4 YIELD STRENGTH AT 0.1 PER CENT OFFSET VERSUS TRUE REDUCTION OF ROD AREA; SAE 1020 STEEL—RODS NOS. 1 AND 2

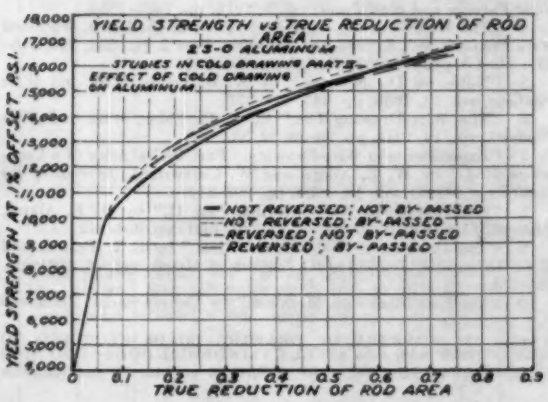

YIELD STRENGTH VS TRUE REDUCTION OF ROD AREA
2S-O ALUMINUM
STUDIES IN COLD DRAWING PART 3 -
EFFECT OF COLD DRAWING ON ALUMINUM.
NOT REVERSED; NOT BY-PASSED
NOT REVERSED; BY-PASSED
REVERSED; NOT BY-PASSED
REVERSED; BY-PASSED
YIELD STRENGTH AT .1% OFFSET P.S.I.
TRUE REDUCTION OF ROD AREA

FIG. 5 YIELD STRENGTH AT 0.1 PER CENT OFFSET VERSUS TRUE REDUCTION OF ROD AREA; 2S-O ALUMINUM

für angewandte Mathematik und Mechanik, vol. 7, 1927, pp. 235–236; "Plasticity Problems in Metals," by George Sachs, Trans. Faraday Society, vol. 24, 1928, pp. 84–92.

11 "Die grundlagen der bildsamen Verformung," by F. Korber and A. Eichinger, Mitteilungen aus dem Kaiser-Wilhelm Institut für Eisenforschung, vol. 22, 1940, p. 57.

12 "Theory of Wire Drawing," by E. A. Davis and S. J. Dokos, *Journal of Applied Mechanics*, Trans. ASME, vol. 66, 1944, pp. A-193–A-198.

13 "An Extension of Wire Drawing Theory With Special Reference to the Contributions of K. B. Lewis," by R. W. Lunt and G. D. S. MacLellan, *Journal of the Institute of Metals*, vol. 72, 1946, pp. 65–96.

14 "A Critical Survey of the Wire-Drawing Theory," by G. D. S. MacLellan, *Journal of the Iron and Steel Institute*, vol. 158, 1948, pp. 347–356. This is an excellent literature review.

15 "The Physical Properties of Steel Wire as Affected by Variations in the Drawing Operations," by H. J. Godfrey, Proceedings of the ASTM, vol. 42, 1942, pp. 513–531.

16 "Plastics for Recording Surface Finish and Wear in Machines," by J. W. Sawyer, Bureau of Ships, Navy Department, presented at the 1949 Seminar on Plastics.

17 Marcokit No. 249 is manufactured by Marco Chemicals, Inc., Sewaren, N. J.

18 "Some Friction Effects in Wire Drawing," by G. D. S. MacLellan, *Journal of the Institute of Metals*, vol. 81, September, 1952–1953, pp. 1–13.

19 Discussion by M. Eugene Merchant of paper, "Sliding Friction Under Extreme Pressures—1," by S. J. Dokos, *Journal of Applied Mechanics*, Trans. ASME, vol. 69, 1947, pp. A-68–A-70.

20 "Sliding Friction Under Extreme Pressures—1," by S. J. Dokos, *Journal of Applied Mechanics*, Trans. ASME, vol. 68, 1946, pp. A-148–A-156.

21 "Spanlose Formung der Metalle," by G. Sachs, *Mitteilungen der Deutschen Material—Prufungsanstalten*, Sonderheit 11–18, 1930–1931, pp. 38–67.

Appendix

Equation [1] is derived by summing up the forces acting on the rod in the longitudinal direction. It is assumed that the longitudinal stresses across any section are uniform, that there is a uniform pressure acting on the rod, and that the average friction coefficient may be used. These assumptions are quite close to the conditions in these tests. The rods entered some dies that had a short entrance section with radial taper before entering the straight-taper section. These entrance conditions require larger

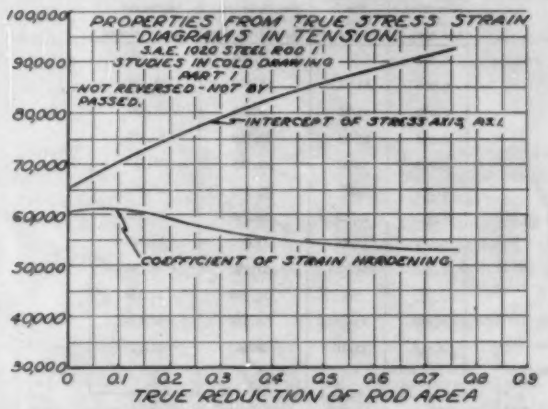

PROPERTIES FROM TRUE STRESS STRAIN DIAGRAMS IN TENSION.
S.A.E. 1020 STEEL ROD 1
STUDIES IN COLD DRAWING PART 1
NOT REVERSED - NOT BY PASSED.
Y-INTERCEPT OF STRESS AXIS P.S.I.
COEFFICIENT OF STRAIN HARDENING
TRUE REDUCTION OF ROD AREA

FIG. 6 PROPERTIES FROM TRUE STRESS-STRAIN DIAGRAM PLOTTED AGAINST TRUE REDUCTION OF ROD AREA. INTERCEPT ON TRUE-STRESS AXIS AND COEFFICIENT OF STRAIN HARDENING; SAE 1020 STEEL—ROD NO. 1

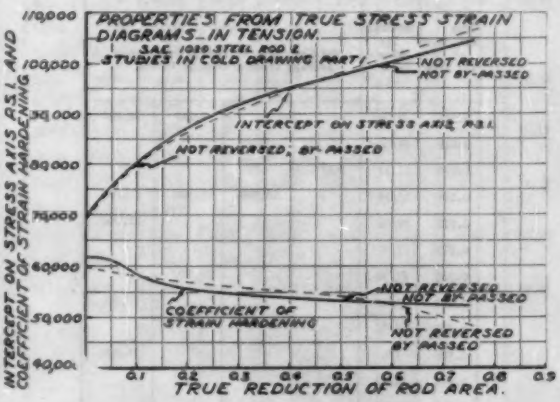

PROPERTIES FROM TRUE STRESS STRAIN DIAGRAMS IN TENSION.
S.A.E. 1020 STEEL ROD 2
STUDIES IN COLD DRAWING PART 1
NOT REVERSED NOT BY-PASSED
INTERCEPT ON STRESS AXIS P.S.I.
NOT REVERSED; BY-PASSED
COEFFICIENT OF STRAIN HARDENING
NOT REVERSED NOT BY-PASSED
NOT REVERSED BY PASSED
INTERCEPT ON STRESS AXIS P.S.I. AND COEFFICIENT OF STRAIN HARDENING
TRUE REDUCTION OF ROD AREA.

FIG. 7 PROPERTIES OF SAE 1020 STEEL, ROD NO. 2 VERSUS TRUE REDUCTION OF ROD AREA

TABLE 16 DETERMINATION OF COEFFICIENT OF FRICTION IN ROD-DRAWING

NOM DIE DIA.	SAE 1020 STEEL			2S-O ALUMINUM		
	DAVIS AND DOKOS EQUATION	SACH'S EQUATION	AVERAGE PRESSURE MEASUREMENTS	DAVIS AND DOKOS EQUATION	SACH'S EQUATION	AVERAGE PRESSURE MEASUREMENTS
1"	.02	.02	.08	.03	.06	.05
63/64	.06	.11	.12	.04	.11	.13
31/32	.09	.15	.16	.06	.10	.11
30/32	.08	.10	.11	.05	.08	.08
29/32	.09	.10	.09	.06	.07	.09
28/32	.08	.09	.14	.05	.07	.13
27/32	.12	.14	.17	.08	.10	.14
26/32	.09	.10	.12	.07	.08	.10
25/32	.11	.12	.22	.06	.08	.16
24/32	.15	.16	.15	.06	.07	.08
23/32	.11	.12	.15	.07	.08	.17
22/32	.13	.14	.20	.07	.09	.15

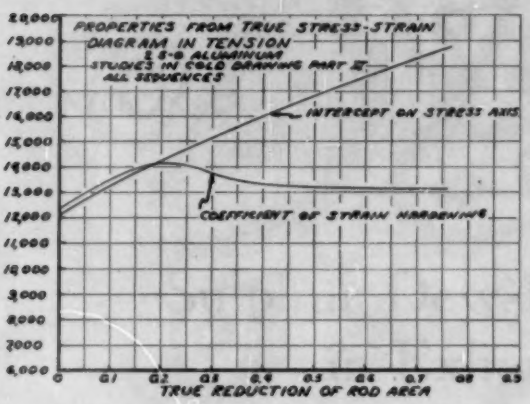

PROPERTIES FROM TRUE STRESS-STRAIN DIAGRAM IN TENSION
2S-O ALUMINUM
STUDIES IN COLD DRAWING PART III
ALL SEQUENCES
INTERCEPT ON STRESS AXIS
COEFFICIENT OF STRAIN HARDENING
TRUE REDUCTION OF ROD AREA

FIG. 8 TRUE STRESS-STRAIN PROPERTIES OF 2S-O ALUMINUM VERSUS TRUE REDUCTION OF ROD AREA

drawing forces than if the entire entrance were in straight-taper sections as reported by Francis and Thompson (9).

Tables 6 to 15 summarize the average drawing loads employing various sequences in cold-drawing.

Fig. 4 illustrates the variation of yield strength at 0.1 per cent offset versus true reduction of rod area for SAE 1020 steel, rods Nos. 1 and 2.

Fig. 5 is a plot of yield strength at 0.1 per cent offset versus true reduction of rod area for 2S-O aluminum.

Figs. 6, 7, and 8 show the properties that are obtained from true stress-strain diagrams. Intercept on the true-stress axis and coefficient of strain hardening are plotted versus true reduction of rod area. Fig. 6 is for SAE 1020 steel rod No. 1, Fig. 7 is for SAE 1020 steel, rod No. 2, and Fig. 8 is for 2S-O aluminum.

Fig. 9 compares Sachs' theoretical equation for rod-drawing with that of Davis and Dokos.

There is tabulated in Table 16 a comparison of the theories of Sachs, and Davis and Dokos with the results from the pressure measurements in this work. For the larger draws where a sufficient cold-working has taken place, there is close agreement.

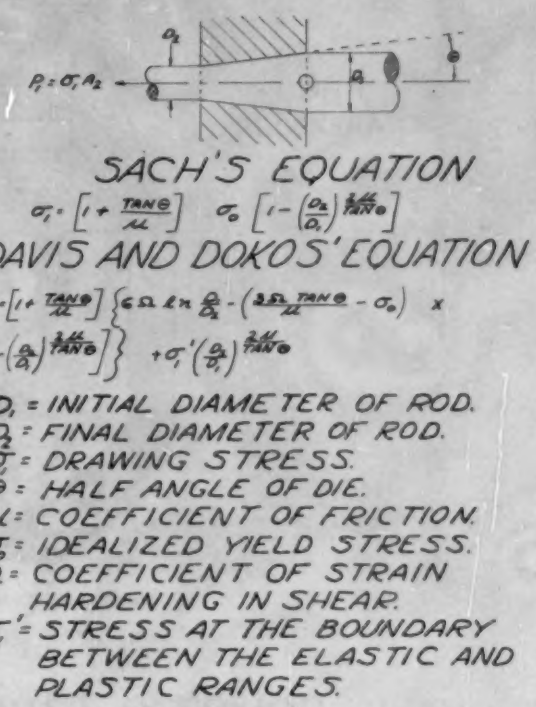

$P_1 = \sigma_1 A_2$
SACH'S EQUATION
$\sigma_1 = \left[1 + \frac{\tan\theta}{\mu}\right] \sigma_o \left[1 - \left(\frac{D_2}{D_1}\right)^{\frac{2\mu}{\tan\theta}}\right]$
DAVIS AND DOKOS' EQUATION
$\sigma_1 = \left[1 + \frac{\tan\theta}{\mu}\right] \left\{ c\Omega \ln \frac{D_1}{D_2} - \left(\frac{2\Omega}{\mu}\tan\theta - \sigma_o\right) \times \left[1 - \left(\frac{D_2}{D_1}\right)^{\frac{2\mu}{\tan\theta}}\right] \right\} + \sigma_1' \left(\frac{D_2}{D_1}\right)^{\frac{2\mu}{\tan\theta}}$
D_1 = INITIAL DIAMETER OF ROD.
D_2 = FINAL DIAMETER OF ROD.
σ_1 = DRAWING STRESS.
θ = HALF ANGLE OF DIE.
μ = COEFFICIENT OF FRICTION.
σ_o = IDEALIZED YIELD STRESS.
Ω = COEFFICIENT OF STRAIN HARDENING IN SHEAR.
σ_1' = STRESS AT THE BOUNDARY BETWEEN THE ELASTIC AND PLASTIC RANGES.

FIG. 9 COMPARISON OF THEORETICAL EQUATIONS—SACHS VERSUS THAT OF DAVIS AND DOKOS

The discrepancy can be explained by the choice of yield strength, especially under a true reduction of rod area of 0.15 where the plastic-hardening region has a variable coefficient of strain hardening. Also, some dies did not offer a complete conical surface for cold-working which was used in the derivation of the theoretical equations.

Discussion

E. S. NACHTMAN.[6] The technique developed by the author for measuring effective die pressures is valuable and essentially simple, his results indicating that the coefficient of friction does not vary appreciably with the die pressure. This agrees with work done by J. G. Wistrech.[7]

The work performed by J. G. Wistrech indicates a coefficient of friction between 0.02 and 0.03 for copper wire, tool steel dies, and sodium stearate. The technique used was the "split" die technique developed by MacLellan.[8] This procedure provides a direct method of measuring frictional effects.

May we hope that the author will continue his work and fit his data to the various wire-drawing theories.

[6] Manager, Product Development Laboratory, La Salle Steel Company, Hammond, Ind. Assoc. Mem. ASME.

[7] "Experimental Investigation of the Mechanics of Wire Drawing," B. I. S. R. A. List No. 60.

[8] Author's bibliography (18).

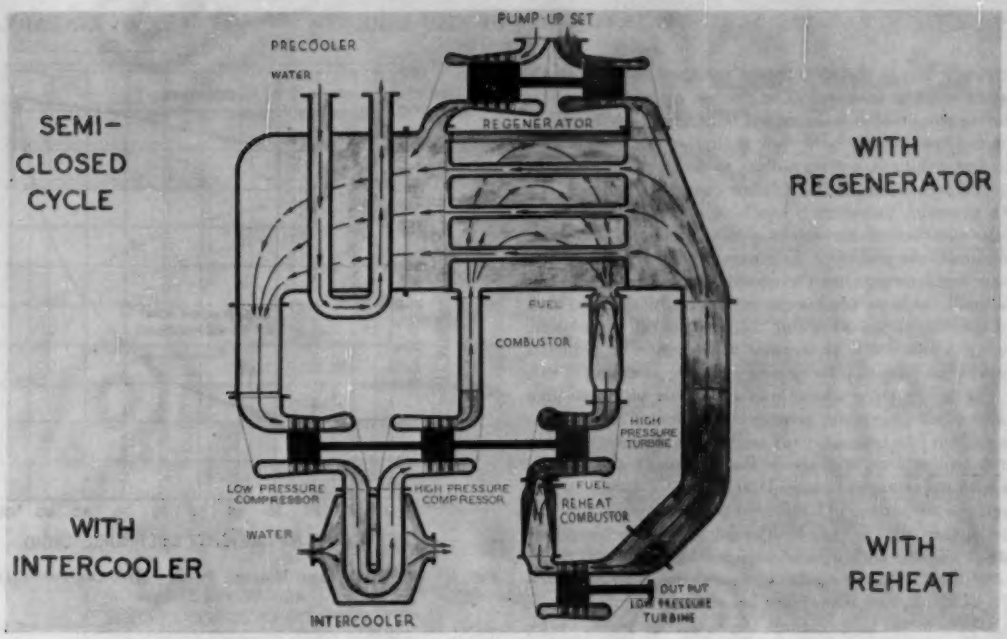


FIG. 1 TYPICAL SEMICLOSED GAS-TURBINE CYCLE

Sulphur Dew-Point Corrosion in Exhaust Gases

BY R. L. COIT,[1] PHILADELPHIA, PA.

Sulphur dew-point corrosion continues to be an important problem in steam-power stations and some gas-turbine applications. A series of corrosion tests is described in which it was found that the effects of this type of corrosion can be minimized by (*a*) use of certain corrosion-resistant alloys such as Hastelloy C, Carpenter No. 20, or Type 316 stainless steel; (*b*) the use of low-sulphur fuels; or (*c*) the maintenance of surface temperatures above the acid-water dew point. The tests indicate that none of the presently available coating materials can provide adequate low-temperature corrosion protection for heat-exchanger surfaces or compressor blading.

INTRODUCTION

IN MANY cases gas-turbine power-plant designers must cope with the same type of low-temperature corrosion that has plagued steam-plant operators for many years.

In the gas-turbine field low-temperature corrosion problems occur at any point where the exhaust gases contact cool surfaces as, for example, in the semiclosed cycle where it is necessary to cool the products of combustion through the dew point, extract the condensed water which was formed during combustion, and then return the gases to the low-pressure compressor. A typical semiclosed cycle is shown in Fig. 1.

Another application in which dew-point corrosion is likely to occur is in waste-heat-recovery devices for gas turbines. Gas-turbine exhaust gases contain large amounts of energy at an initial temperature level of 700 to 800 F. These exhaust gases can be utilized to heat feedwater, to heat sea water for use in a flash evaporator which converts sea water into potable water, or to provide preheated air for a conventional steam boiler. In each of these applications the exhaust gases contact cool surfaces and the danger of dew-point corrosion exists.

EFFECT OF SULPHURIC ACID UPON DEW POINT OF FLUE GASES

The flue gases from steam generators or gas turbines usually have a calculated water dew point of the order of 100 F. The surfaces of the various heat-recovery components are normally arranged so that the metal temperatures are well above the calculated dew point of the exhaust gases, and in this way condensation of moisture is prevented.

In many instances, however, it is found that part of the sul-

[1] Development Engineer, Steam Division, Westinghouse Electric Corporation. Assoc. Mem. ASME.

Contributed by the Gas Turbine Power Division and presented at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–21, 1955, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 3, 1955. Paper No. 55—S-34.

phur in the fuel leaves the combustion zone not as sulphur dioxide, but as sulphur trioxide. The process by which the SO_2 formed during combustion is converted to SO_3 is not well understood, but the most widely held view is that substances such as iron oxide, iron sulphate, and vanadium oxide act as catalysts to accelerate the conversion of the sulphur dioxide to sulphur trioxide. An extensive literature is available on the rates of reaction and the equilibrium conversion of SO_2 to SO_3, but catalytic effects complicate the problem. In general, the reaction rates are high at high temperatures but the equilibrium conversion to the trioxide is small; at lower temperatures the possibility exists that additional conversion can occur but the reaction rates are lower. The action of a catalyst is to increase the reaction rates in the lower-temperature regions thus causing greater conversion of SO_2 to SO_3. The use of large amounts of excess air in gas-turbine practice also tends to produce greater conversion of SO_2 to SO_3. After the sulphur has been oxidized to SO_2, a gas-phase reaction occurs at a temperature well above the dew point and the SO_3 combines with water vapor to form H_2SO_4.

Although the percentage of sulphur in the fuel is small and the proportion of this sulphur that is oxidized to sulphur trioxide is also small, it may be sufficient to increase the dew point by as much as 250 deg F. The dew point of the exhaust gases may then be as high as 350 F, and it may not be economically feasible to maintain all the metal temperatures above the dew point. Condensation of sulphuric acid occurs on these cool surfaces and intensive corrosion results.

Early Studies. Johnstone (1)[2] first emphasized the marked effect that small concentrations of sulphuric acid could have on the dew point of flue gases. This work was later extended to include lower-acid concentrations by Taylor (2). Thomas and Barker (3) measured the partial pressures of water vapor and sulphuric-acid vapor, and these data can be used to construct boiling-point-composition diagrams for any pressure. The influence of sulphuric acid upon the dew point of gases containing different concentrations of water vapor is shown in Fig. 2. It is seen that the dew point is mainly dependent upon the amount of acid present and depends only slightly upon the concentration of water vapor.

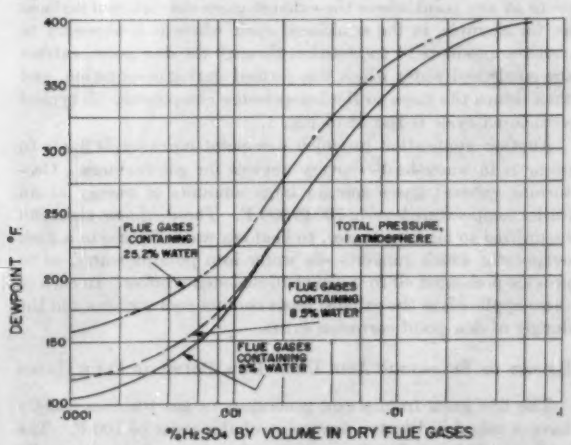


FIG. 2 EFFECT OF SULPHURIC ACID ON DEW POINT

Fig. 3 is a typical boiling-point composition diagram for water-sulphuric acid mixtures. If the concentrations of water vapor and sulphuric-acid vapor are known, the temperature at which the first droplet of liquid is formed when the gases are cooled can be found. The dotted line shows what happens when gases containing 88 per cent by weight water vapor and 12 per cent sulphuric acid are cooled. The acid content of 12 per cent was chosen for illustration of the two-phase system and is not representative of acid contents to be found in gas-turbine practice. At 419 F the vapors will be in equilibrium with a liquid of 94.5 per cent H_2SO_4. If in a closed system the vapors are further cooled, the dew-point temperature will decrease as condensing continues, since the condensed liquid is more concentrated in acid than is the vapor. At the same time the condensing liquid will become less and less concentrated. In the end, if the temperature is sufficiently low, the concentration of all the liquid condensed will be that of the original vapor, i.e., 12 per cent H_2SO_4. Fig. 3 was calculated for a total pressure of condensable vapors of 50 mm Hg.

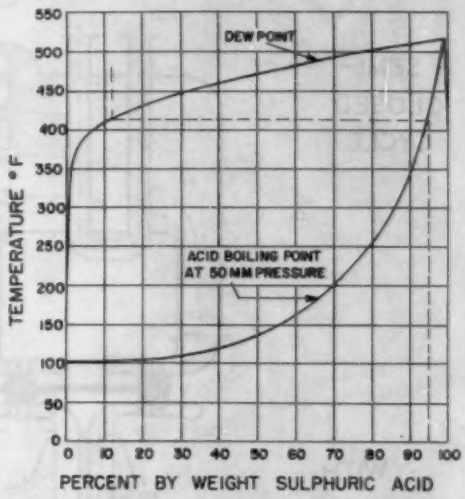


FIG. 3 DEW POINT AND BOILING POINT—DIAGRAM FOR SULPHURIC ACID-WATER SYSTEM

Calculating Strength of Sulphuric-Acid Deposits. In order to estimate the maximum strength of the sulphuric acid which is deposited on cooled surfaces one must know the sulphur content of the fuel oil, the percentage of the available sulphur which is oxidized to SO_3, and the gas-fuel ratio in the system. Fig. 4 was calculated by using typical gas-fuel ratios for gas turbines, and assuming various sulphur contents in the fuel and a range of percentage conversions to SO_3. From this curve it is noted that for a sulphur content of 1.0 per cent and complete conversion of the sulphur to SO_3, the dew point is approximately 345 F. The acid concentration corresponding to this dew point is found from Fig. 5 to be approximately 89 per cent. This information concerning the percentage sulphur converted to SO_3 and the maximum acid concentration to be expected could be determined simply by knowing the sulphur content of the fuel and making a dew-point observation with the B.C.U.R.A. dew-point meter (4). For example, if the fuel contains 0.5 per cent sulphur and the dew point is measured to be 270 F, we find from Figs. 4 and 5 that 50 per cent of the available sulphur has been converted to SO_3 and the maximum acid concentration is 80 per cent.

In the foregoing example we determined the maximum acid concentration to be expected. Another approach is to assume that the condensing occurs at the temperature of the cooled surface rather than at the dew point of the gases. Taylor (2) measured the acid concentration in the condensate from gases of vary-

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

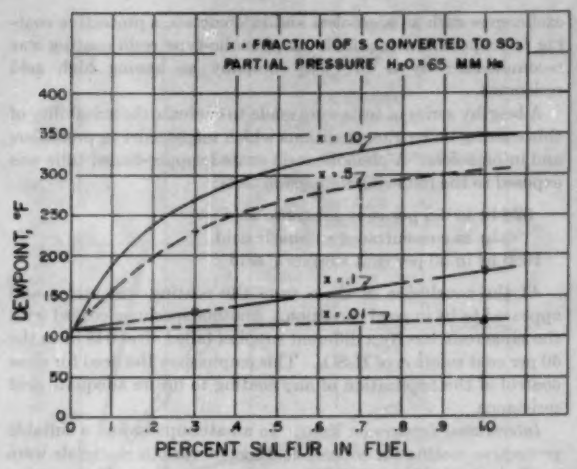


FIG. 4 EFFECT OF SULPHUR ON DEW POINT

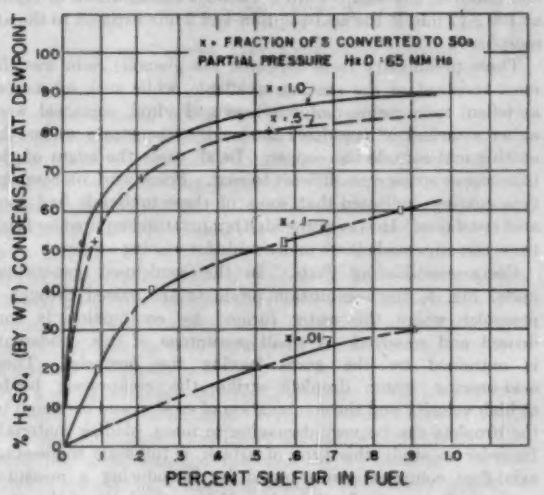

x = FRACTION OF S CONVERTED TO SO3
PARTIAL PRESSURE H2O = 65 MM Hg
x = 1.0
x = .5
x = .1
x = .01
% H2SO4 (BY WT) CONDENSATE AT DEWPOINT
PERCENT SULFUR IN FUEL

FIG. 5 EFFECT OF SULPHUR ON ACID CONCENTRATION

ng composition. It was found that for gases containing 0.007 to 0.07 per cent sulphuric acid by volume, the acid concentration of the condensate which is obtained in a constant temperature condenser is independent of the original gas composition. For a condenser temperature of 243 F, the data given in Table 1 were

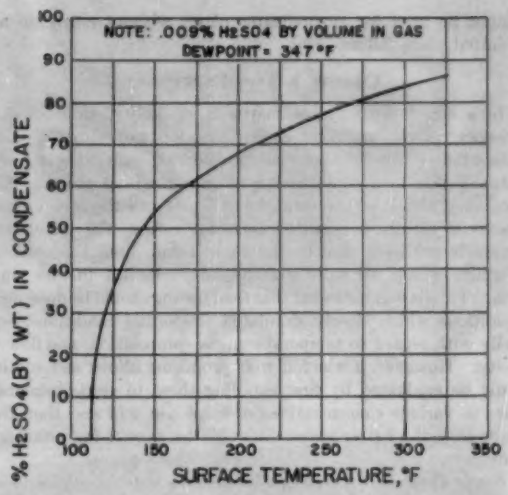


FIG. 6 VARIATION OF CONDENSATE STRENGTH WITH TEMPERATURE

the concentration of the acid on the surface depends only on the temperature of the surface and is independent of the initial H_2SO_4 content of the gases.

Fig. 6 shows the marked effect of the surface temperature upon the acid concentration of the condensate. Since in a practical application the surface temperatures of various areas of the air heater, feedwater heater, precooler, or other device will be different, the materials chosen must be resistant to acids of a wide range of concentration.

PREVIOUS INVESTIGATIONS

Much of the early work on dew-point corrosion was done under the sponsorship of the British Boiler Availability Committee. The results of laboratory tests and of field tests made on stoker-fired steam plants have been reported in the literature (4). It was found that the introduction of substances such as pulverized-coal, fly ash, carbon smoke, zinc, and magnetite (Fe_3O_4) into the combustion zone had a beneficial effect on flue gases which contained appreciable quantities of sulphur trioxide.

Extensive evaluation of the effect of additives was made at several oil-fired steam stations in this country (5). These investigations indicate that the use of an additive such as dolomite (a natural mineral composed of approximately 60 per cent calcium carbonate and 40 per cent magnesium carbonate) will reduce the amount of sulphur trioxide in the flue gases and will inhibit corrosion.

A comprehensive study of low-temperature corrosion in regenerative air preheaters (6) indicated that Hastelloy B and C,

TABLE 1 ACID CONCENTRATIONS FOR CONDENSER TEMPERATURE OF 243 F

% H_2SO by volume	Dew Point, deg F	% H_2O by volume	Measured acid concentration in condensate, wt per cent	Calculated acid concentration in condensate, wt per cent	Calculated acid concentration at dew point, wt per cent
0.007	320	8.3	76.0	75.3	86
0.009	347	8.3	75.5	75.3	88
0.025	374	8.2	75.3	75.4	92
0.07	401	8.5	75.3	75.1	94

obtained. Other workers (7) have shown that the composition of the condensate depends upon the catalytic activity and chemical nature of the condensing surface as well as its temperature.

This indicated that when acid is condensing on a cool surface

Inconel and Carpenter No. 20 had lower rates of corrosion than Corten (a low-alloy steel) and that all high-alloy steels, with the exception of Carpenter No. 20 had higher corrosion rates than Corten. None of the 22 coated or plated materials resisted corrosion satisfactorily. In general, metals that form sulphates

cannot be used for air-preheater plates without corrosion and eventual plate failure.

Corrosion-Test Experience

In a search for suitable materials of construction for compressors, valves, and heat exchangers, it was realized that the deleterious effects of these sulphur-bearing exhaust gases were actually due to a combination of corrosion and erosion. The combined action of the condensing liquid and the flow of gases results in greatly accelerated corrosion rates. This accelerated action is probably due to the rapid removal of the corrosion products which normally would tend to inhibit further corrosion. Thus it was indicated that final testing should be done under conditions which closely simulated the actual conditions, especially with regard to temperature, gas composition, and flow velocity. However, it was felt that promising alloys and coatings could be evaluated by first exposing them to static immersion tests in various concentrations of sulphuric acid and then from these tests the better materials would be chosen for rotating or flow tests.

Static Tests. In a series of immersion tests in sulphuric-acid concentrations of 0.01, 0.03, and 0.8 per cent, it was found that

TABLE 2 IMMERSION TEST RESULTS IN 50 PER CENT SULPHURIC ACID AT 150 F

Material	Hours	Weight loss, per cent	Remarks
N155	4.5	26	Rapid attack
25/20	4.5	39	Rapid attack
18/8	4.5	22	Rapid attack
Above specimens removed after 4.5 hr due to rapid attack			
Titanium	6	1.3	Active attack
12% chrome steel	6	2	Active attack
Nickel-plated	53	0.65	Failed after 53 hr
12% chrome steel	53		
Inconel	360	7.4	
Al-bronze plate	360	0.50	
Al-bronze plate	1080	0.84	
American brass	360	0.33	
70–30 Cu-Ni	1080	0.78	
Revere Al-bronze	360	0.50	
Revere Al-bronze	1080	0.78	
Ambraloy No. 901	360	0.45	
Ambraloy No. 901	1080	0.74	
Hastelloy C	360	0.04	
Hastelloy C	1080	0.18	
Carpenter No. 20	360	0.14	
Carpenter No. 20	1080	1.07	

N155, 25/20 and 18/8 stainless steels, Refractaloy and Discaloy, had good acid resistance while copper, Inconel, 70–30 Cu-Ni, and Admiralty metal were not satisfactory. The compositions of the alloys tested are listed in Table 3.

The results of additional tests made in 50 per cent H_2SO_4 at 150 F (150 F for 8 hr a day with heat off nights and week ends) are given in Table 2.

exchangers such as precoolers and intercoolers, a protective coating for copper was required. A phenolic-type resin coating was recommended by a chemical company as having high acid resistance.

A lengthy series of tests were made to evaluate the suitability of this coating under the conditions which might exist in precoolers and intercoolers. A phenolic resin-coated copper-finned tube was exposed to the following immersion tests:

528 hr in 0.8 per cent sulphuric acid
96 hr in concentrated sulphuric acid
1080 hr in 50 per cent sulphuric acid

At the conclusion of these tests the coating was intact and appeared to be in good condition. Another specimen coated with the same resin but by a different supplier failed after 648 hr in the 50 per cent solution of H_2SO_4. This emphasizes the need for close control of the application of any coating to insure adequate acid resistance.

Intermittent Immersion Tests. In an attempt to find a suitable protective coating for copper, coatings of various materials were applied to thin (0.005–0.015 in.) copper strips. These strips were then attached to a motor-driven rig which periodically dipped and removed the samples from various concentrations of H_2SO_4 at 100 F (2 min in the acid solution and 2 min exposed to the atmosphere).

These preliminary tests indicated the phenolic resin was the most resistant of the coatings available, while such substances as teflon, epon resins, polyethylene and vinyl organosal were either attacked or permitted the acid to penetrate under the coating and corrode the copper. In all cases the edges of the thin copper strips were difficult to coat. Some work on ceramic-type coatings indicated that some of these materials had good acid resistance. However, the high temperature required for firing these ceramics made them undesirable for coating copper.

Compressor-Blading Tests. In the semiclosed gas-turbine cycle, Fig. 1, the combustion products are passed through a precooler where the water formed by combustion is condensed and removed. A small percentage of this condensate is entrained by the gases leaving the precooler. These acid-bearing water droplets strike the compressor blades at high velocity and the combination of erosion and corrosion by the droplets can be very damaging to many blading materials. In order to study this type of attack, a full-scale single-stage axial-flow compressor was tested by introducing a measured mixture of steam and sulphuric acid into the inlet of the compressor. The acid content of the water-acid mixture was maintained at 1 per cent by weight. The compressor blades were manufactured from a total of ten materials and five different

TABLE 3 TYPICAL COMPOSITIONS

No.	Material	C	Mn	Cr	Ni	Cu	Fe	Other
1	Aluminum Bronze					95		5 Al
2	Al-Ni Bronze				5	82	2.5	9.5 Al
3	Carpenter No. 20	0.07	0.75	20	29	3	44.2	2 Mo
4	Corten	0.1	0.4	0.9	0.45	0.4	97	
5	Discaloy			13.5	26		54	1.6 Ti
6	Hastelloy B	0.1	1.0	1.0	63.5			28 Mo
7	Hastelloy C	0.15	1.0	16.5	54			17 Mo, 4W
8	Inconel	0.08	0.25	14	78.7		6.5	
9	Mayari-R	0.1	0.65	0.6	0.4			
10	12 per cent chrome	0.04	0.5	12.7	0.12		86.2	
11	70–30 copper-nickel				30	69		1 Sn
12	Type 304 stainless	0.05	1.5	19	10		68	
13	Type 310 stainless	0.2	1.5	25	20		52	
14	Type 316 stainless	0.05	1.5	19	12		65	

It is seen that some materials which were satisfactory at low acid concentrations were rapidly attacked by 50 per cent acid. Hastelloy C appeared to be the most resistant of the metals with 70–30 copper-nickel having good acid resistance also.

Since it is often desired to use copper fins on the gas side in heat

protective coatings were applied to the various blade materials. The relative weight loss of the alloys as a function of the testing time is shown in Fig. 7. The Corten, Mayari-R, 12 per cent chrome, 12 per cent chrome with nickel plate, aluminum-bronze and aluminum-nickel-bronze blades showed high rates of corro-

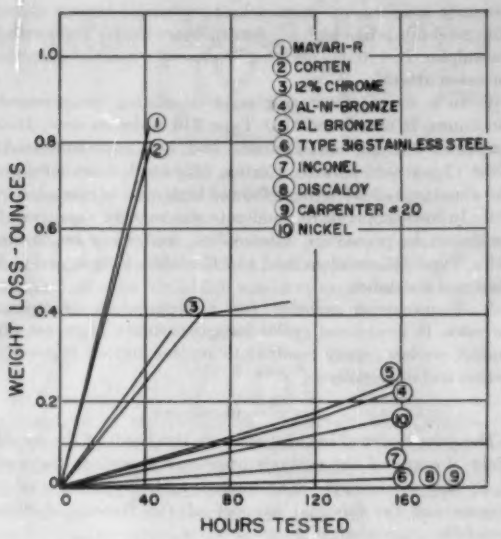


FIG. 7 EFFECT OF TESTING TIME ON WEIGHT LOSS OF ALLOYS

sion and/or erosion, while the best materials were Inconel, Carpenter No. 20, Type 316 stainless steel, and Discaloy.

None of the commercially available protective coatings was satisfactory for protecting compressor blades, but it appeared that coatings could withstand the action of the gases in regions of lower velocity, such as exist in the inlet casing and the inlet guide vanes.

Model Heat-Exchanger Tests. Two model heat exchangers were built and tested to determine the relative corrosion resistance of various alloys and coatings under conditions similar to those found in precoolers or intercoolers, that is, moderate gas velocities and low temperatures.

Fig. 8 shows the finned-tube heat exchanger before it was placed in operation. The finned tubes have a 3/8-in-diam base tube with 0.9 in. over-all diameter and nine 0.012-in-thick fins per in. The copper fins were attached mechanically by cutting a helical groove in the base tube, inserting the fin, and then knurling at the base of the fin. Two of the tubes were coated with a phenolic resin and two were uncoated. Products of combustion at 600 F and 50 psia enter at the top of the exchanger and flow across the finned tubes with an average velocity of about 30 fps. The gases are cooled to 80 F in this exchanger; thus the dew-point type of corrosion would be expected. After some 325 hr of operation with a fuel containing 1 per cent sulphur, it was noted that the phenolic-resin coating was being lifted from the fins in the region where initial condensation occurs. At this time the uncoated fins had been reduced in thickness from 0.012 to 0.005 in. It was found that the corrosive attack was more severe in the regions where the initial condensation occurred. The cooler surfaces downstream from the initial condensation regions were attacked, but in accordance with the theory, the corrosion was less rapid.

The effect of the resin coating on the heat-transfer capability of the finned tubes was evaluated in a separate facility. It was found that the coated tubes would transfer only 80 per cent as much heat as would the uncoated tubes under identical flow conditions.

The second model heat exchanger, Fig. 9, was used to evaluate the relative resistance of several of the more promising alloys. The tubes were 3/8 in. OD and the casing and baffles were made from Type 304 stainless steel. After 531 hr operation with a diesel oil containing 1 per cent sulphur, the following results were observed:

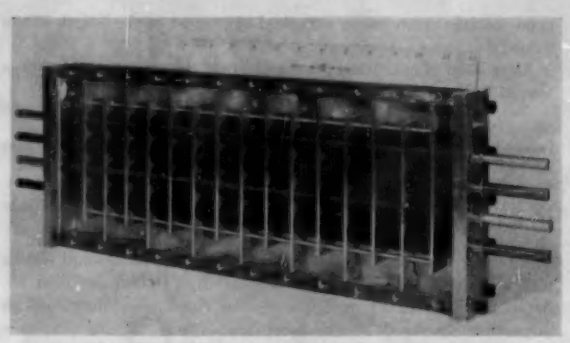


FIG. 8 FINNED-TUBE TEST HEAT EXCHANGER

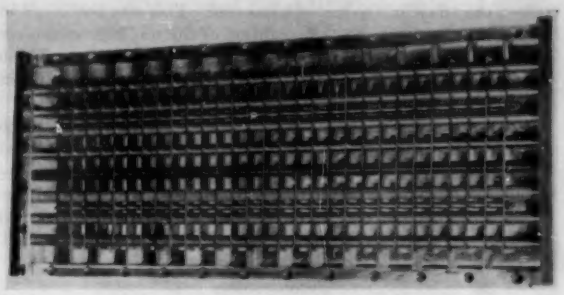


FIG. 9 BARE-TUBE HEAT EXCHANGER

Material	Condition
Type 304, stainless steel	Excellent condition, no attack
Type 316, stainless steel	Excellent condition, no attack
Type 310, stainless steel	Excellent condition, no attack
Inconel	Many pits, principally near hot end
Hastelloy C	Many pits, principally near hot end
Aluminum bronze (No. 911)	Badly corroded over entire length
70–30 copper-nickel	Fair condition, rough, spongy appearance
Carpenter No. 20	Good condition but with many very small pits over entire length

Semiclosed-Cycle Tests. In a further attempt to provide realistic test conditions for the evaluation of the resistance of alloys and protective coatings, a large-scale semiclosed-cycle test facility was set up by the USN Engineering Experiment Station at Annapolis. It was thought that there might be a cumulative effect in the semiclosed cycle, since additional sulphur oxides were continually being introduced into the system from the fuel. A portion of the recirculating gases was passed through a test heat exchanger in which were installed both finned tubes and bare tubes. The gases enter the test heat exchanger at 500 F and 55 psia and are cooled to 110 F. After 50 hr of operation with 0.9 per cent sulphur fuel, it was found that the uncoated copper-finned tubes were badly corroded and were loosened from the base tube. The coated tubes (phenolic resins and ceramic-

type coatings) were badly attacked over most of the length of the tubes.

After 50 hr of operation the condition of the bare-tube materials was as follows:

90–10 Cu-Ni	Etched surface
Inconel	Small pits in condensing region
70–30 Cu-Ni	Etched in condensing region
Type 316 stainless	Good condition
Carpenter No. 20	Pitted
Hastelloy C	Good condition

After the installation of new finned tubes, an additional 150-hr test was made using very low-sulphur fuel (0.005 per cent S). After the first 12 hr of operation on the low-sulphur fuel an inspection disclosed severe corrosion of one of the two new phenolic resin-coated tubes. It appeared likely that the corrosion was a carry-over from the system when it had been operated on high-sulphur fuel, and that the indicated corrosion did not truly reflect the effect of the low-sulphur fuel. This was borne out in subsequent testing, with the condition of the test heat exchanger remaining unchanged through the next 150 hr of testing. These tests indicate that one way to minimize the corrosion problem in a semiclosed cycle is to use a low-sulphur fuel.

Spray Cooling. An experimental study is now being made to evaluate the possibility of replacing the conventionally designed precooler with a spray cooler in which the gases are cooled by contact cooling with water droplets. In this arrangement the condensation of the acid would occur on the water droplets and the acid concentration would be reduced to a less corrosive level.

Conclusions

1 Static exposure tests indicated that while N155, Type 310, and Type 304 stainless steels were satisfactory for use in sulphuric-acid concentrations of less than 0.8 per cent, they suffered rapid attack when exposed to a 50 per cent solution of H_2SO_4. In these static tests, Hastelloy C appeared to be the most resistant of the alloys tested.

2 A phenolic resin used in industry to coat acid containers appeared to be the most suitable coating to protect copper from acid attack. However, this coating was badly attacked in the semiclosed-cycle test heat exchanger when 1 per cent sulphur fuel was used. Thus it is indicated that no suitable coating is presently available to protect heat-exchanger surface materials from corrosion when high-sulphur fuels are used. Tests with very low-sulphur fuel (0.005 per cent S) showed a marked reduction in corrosion attack.

3 In a series of rotating tests simulating compressor-inlet conditions, it was found that Type 316 stainless steel, Inconel, Carpenter No. 20, and Discaloy had high corrosion resistance while 12 per cent chrome, Corten, Mayari-R, aluminum-bronze and aluminum-nickel-bronze showed high rates of corrosion.

4 In low-temperature, moderate gas-velocity regions such as are found in precoolers, intercoolers, and waste-heat-recovery units, Type 316 stainless steel and Hastelloy C appeared to have good acid resistance.

5 A promising scheme for minimizing low-temperature corrosion in semiclosed cycles for gas turbines is to use direct-contact coolers (spray cooling) to replace surface types of precoolers and intercoolers.

Acknowledgments

The data reported in this paper are the result of the combined effort of many of the author's associates as well as the work of U. S. Navy personnel. The interest and co-operation of these persons and the financial support of the Bureau of Ships is gratefully acknowledged.

BIBLIOGRAPHY

1 "Effect of Sulphuric Acid on the Dewpoint of Flue Gases," by H. F. Johnstone, Circular No. 20, University of Illinois Experiment Station, 1929.

2 "Relation Between Dewpoint and Concentration of Sulphur Acid in Flue Gases," by A. A. Taylor, *Journal of the Institute of Fuel*, vol. 16, October, 1942, pp. 25–28.

3 "The Partial Pressures of Water Vapor and of Sulphuric Acid Over Concentrated Solutions of Sulphuric Acid of High Temperatures," by J. F. Thomas and W. F. Barker, *Journal of the Chemical Society*, vol. 127, 1925, p. 2820.

4 "Developments in the B.C.U.R.A. Dewpoint Meter," by J. Corbett, D. Flint, and R. F. Littlejohn, *Journal of the Institute of Fuel*, vol. 25, November, 1952, p. 246.

5 "The Use of Additives for the Prevention of Low-Temperature Corrosion in Oil-Fired Steam-Generating Units," by E. C. Huge and E. C. Piotter, Trans. ASME, vol. 77, 1955, pp. 267–278.

6 "Corrosion and Deposits in Regenerative Air Preheaters," by J. F. Barkley, et al., Bureau of Mines Report of Investigations No. 4996, August, 1953.

7 AGA Proceedings, vol. 17, 1935, pp. 706–725.

Water Tunnels for Hydraulic Investigations

By J. M. ROBERTSON,[1] URBANA, ILL.

The water tunnel is defined as a device similar to a wind tunnel for cavitation and other hydraulic testing. A list is presented of tunnels in the United States, and the various types of tunnels are described. Cavitation, the principal phenomenon of interest in most water-tunnel researches, is discussed and its scaling factor defined. The purposes and uses for which water tunnels have been built or employed are reviewed. A bibliography on cavitation and water tunnels is included.

Introduction

THE material included in this paper had its origin in a talk prepared at the request of the Hydraulics Division for general presentation at a session of the Engineering Centennial Convocation in Chicago, Ill., September, 1952. As more material was collected than could be included in the talk, an extended version was issued as an external report of the Ordnance Research Laboratory.[2] With inclusion of material which has become available since then, the present paper presents a summary of the study for the benefit of those to whom the report is not available. In this paper emphasis is placed on explaining to the uninitiated what a water tunnel is and in describing the salient features of such a facility, as well as in discussing uses for water tunnels. For brevity's sake, material on tunnel-design features has been omitted—such material is contained in the reference report.[2] However, there is included here a 60-item Bibliography on cavitation and water tunnels.

It is hoped that the information presented will be of assistance and encouragement to those interested in the use of water tunnels, in general, hydraulic research. If this presentation does nothing else it certainly reveals the many uses and potentialities of this type of research facility.

Definition of a Water Tunnel

In this paper a water tunnel is defined as the hydraulic analog of the aerodynamicist's air tunnel. To almost everybody it is clear what the air or wind tunnel used in aerodynamic testing is. With an air tunnel, the relative flow of air around an object is studied in the laboratory. Thus the flow around an airplane wing, such as occurs when the airplane moves through the air, is studied in the tunnel with the wing, or model of it, held fixed and the air moved past it. In water-tunnel testing the same approach is used. A water tunnel, therefore, is a research or test facility in which the flow about bodies, which move in water, is studied by holding them fixed and moving the water past them. For cavitation studies in a water tunnel, one must be able to vary the pressure in the test region. Hence the ability to vary the pressure is as important in a water tunnel as is the ability to vary the speed.

The first research facility resembling a water tunnel was built some 60 years ago in England by C. A. Parsons for propeller studies. This was rather small (over-all length about 3 ft), and the test propeller produced the water circulation. A second propeller tunnel was built in Germany at about the same time. In the decade preceding the second world war, some ten tunnels were built. Most of these were propeller tunnels for use by naval architects although a few were for other purposes.

If one considers water tunnels as covered in this paper and restricts his counting to those of a size large enough to permit adequate testing of bodies and forms in them, then one can count over a dozen in the United States alone. These water tunnels of 6-in. throat size and greater are listed in Table 1, with an indication of their location and principal features. The test speed available is seen to vary from 20 to 100 fps and the power required from 15 to 2000 hp. The significance of the various types of jet and circuit will be indicated later on in this paper.

If this table were extended to cover the Western Hemisphere, only a few more tunnels would need to be added. Europe is well equipped with water tunnels. England has at least four in operation, and as many more, some of which are relatively large, in various stages of design and construction. Other European water tunnels are in France, Holland, Sweden, and Spain. Russia certainly has more than the two tunnels referenced in the literature. A large number of tunnels were constructed in Germany but these were removed after the war; Japan has at least two tunnels. Water tunnels, therefore, are not too uncommon.

Cavitation

As water tunnels are designed primarily for hydraulic studies involving cavitation, a discussion of this phenomenon is appropriate. Cavitation is the formation of holes or voids in water due to the occurrence of low pressures. This boilinglike phenomenon is usually found to occur whenever the absolute pressure at a point in a liquid stream is reduced to the vapor pressure. The cavities formed, therefore, are filled with vapor and any gases dissolved in the water.

Cavitation produces three undesirable effects which should be avoided in hydraulic devices. With the first occurrence of cavitation, noise is produced. In naval work this is undesirable because of the use of acoustic underwater listening devices and weapons. As the amount of cavitation increases on the object a second effect, a loss of efficiency or effectiveness, occurs. This of course is undesirable. The third trouble is found when excessive cavitation is allowed to occur for some time. This produces cavitation erosion in which portions of the structure involved are eaten away. The last two effects are causes for worry to designers of dams and conduits, turbines, pumps, ship propellers, etc.

Cavitation Index. In studying the cavitation phenomena one needs a scaling parameter similar to the Reynolds number or to the Mach number used by aerodynamicists for compressibility effects. Although there is not complete uniformity in the use of the term, most cavitation scaling may be done with the aid of an index or number for which the letter sigma is commonly used. The original suggestion for this index probably came from Thoma. This index or number is the ratio of the available static pressure (above vapor pressure) to the dynamic pressure of the oncoming flow, or simply

[1] Professor of Theoretical and Applied Mechanics, College of Engineering, University of Illinois; formerly, Professor of Engineering Research, Ordnance Research Laboratory, The Pennsylvania State University, University Park, Pa. Mem. ASME.

[2] Ordnance Research Laboratory, The Pennsylvania State University, External Report No. NOrd 7958-256, March 31, 1953.

Contributed by the Hydraulic Division and presented at a joint session with the Applied Mechanics Division at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–21, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 20, 1955. Paper No. 55—S-19.

TABLE 1 WATER TUNNELS IN THE UNITED STATES

Date	Location	Size, in.	Max speed, fps	Motor, hp	Type		Use
					Jet	Circuit	
1929	David Taylor Model Basin (USN)	12 (diam)	26	15	Open	Closed	Propeller tests
1938	Mass. Inst. of Tech.	20 (diam)	34	75	Open	Closed	Propeller tests
1940	David Taylor Model Basin (USN)	24 (diam)	60	400	Open	Closed	Propeller tests
1942 (47)	Calif. Inst. of Tech.	14 (diam)	100	350	Closed	Closed	Body studies; resorber
1944 (1951)	State Univ. of Iowa now at U. S. Waterways Expt. Sta.	13 (diam) 12 × 12	35	80	Open and closed Closed	Closed	Body studies, etc.
1946	The Penna. State University (dismantled)	6 (diam)	50	170	Closed	Open	Experimental
1946	Calif. Inst. of Tech.	20 × 21	30	75	Free surface	Closed	Air separator
1947 et seq.	Univ. of Minnesota	6 (diam)	50	15	Closed and open	Closed	Model
1948	Newport News Shipbldg. and Dry Dock Co. (On loan to Reed Res. Co.)	7 × 24 (2½ × 20) 6 × 24	21 (60)	60	Free surf. and clo.	Closed	Model and test (Skin friction; flex. walls)
1949	State Univ. of Iowa	13 (diam) 12 × 12	35	40	Open and closed	Closed	Body studies, etc.
1950	ORL—The Penna. State University	48 (diam)	83	2000	Closed	Closed	Body-propeller systems
1951	Illinois Inst. of Tech. (dismantled)	12 (diam)	...	...	Closed	Open	Drag tests (no pressure control)
1951	Calif. Inst. of Tech.	12 × 12	15	(low)	Closed	Closed	Turb. diff. studies (no pressure control)
1951	University of Minnesota	10 (diam)	50	...	Free	...	Free jet; gravity flow
1952	ORL—The Penna. State University	12 (diam) 4½ × 20 6 × 24	70	150	Closed and open	Closed	Body studies, etc.
1952	State University of Iowa	13 (diam) 12 × 12	35	40	Open and closed	Closed	Body studies, etc.

$$\sigma = \frac{p - p_v}{\frac{1}{2}\rho V^2}$$

where p is the absolute hydrostatic pressure in the undisturbed flow and p_v the vapor pressure, V the velocity of the relative flow, and ρ the liquid density. In turbine and pump work this index is sometimes defined a little differently, often in terms of the gross-flow characteristics of the unit. However, in analyzing the cavitation characteristics of the flow about an impeller it will be found that the best correlation with hydrofoil characteristics is obtained if the foregoing type of index is used in terms of the local conditions at the section involved.

In discussing cavitating conditions on a body it is not sufficient to state the operating cavitation index. One also must know the critical index of the body for inception of cavitation. The lower the operating index the more prone the body is to cavitate or, if the inception point has already been passed, the more severe will be the cavitation. Thus, if one has a body with a critical cavitation index of 1.0 operating at an index of 1.5 or even 1.05, there is no cavitation. However, if the operating index is 0.99 the body is cavitating and if the operating index is lower, say, 0.9, it is undergoing severe cavitation. For cavitation there is not a unique critical value of the parameter dividing the region of cavitation from that of no cavitation, but rather there is a different critical index for each shape or body. However, critical indexes for bodies are of the same order of magnitude; thus a well-designed body may have a critical index of 0.2 and a very blunt shape may have an index as high as 2.0.

Cavitation–Compressibility Analogy. Various people have noted an analogy between some of the features of cavitation in liquids and compressibility in gases. Thus, in both cases, trouble (cavitation or shock waves) first occurs at the point along the surface where the velocity is highest. The body or foil, most suitable to delay as long as possible the effect in either case, is one of constant velocity along the surface. During the war Reichardt in Germany used water-tunnel cavitation studies to develop shapes for high-velocity (near or transonic) aerodynamic parts. Another application of this analogy involves the use of water-tunnel tests to determine the locations where shock waves will first occur on high-speed bodies or combinations of bodies.

Tunnel Purposes or Uses

Water tunnels, like other research facilities, are fairly flexible as to their type of application. However, they will be classified according to the purposes or uses for which they were primarily built. Thus we have propeller tunnels, tunnels built for studying the forces and cavitation around body shapes, experimental or model tunnels built to improve our knowledge of flow in tunnels and hydraulic components, etc. Initially, most of the research or testing tunnels were built for propeller work; however, the number of those being constructed for other purposes is increasing. Reference to Table 1 shows that most tunnels in the United States are nonpropeller ones. In Europe the reverse is true. Some aerodynamic uses for water tunnels have been noted elsewhere (D20, D22).[3]

Propeller Tunnels. In water-tunnel propeller testing almost all studies are made in what is termed the free-stream condition. The propeller is held in the test or working section of the tunnel and is rotated at the proper speed as the tunnel water is forced past it. It is usually operated at, or near, the correct advance coefficient (ratio of the water speed to the product of the rational speed and propeller diameter). Cavitating conditions are simulated by reducing the pressure in the working section. Fig. 1 is a general view of a propeller under test in the 12-in. water tunnel of the U. S. Navy's David Taylor Model Basin. This picture gives an idea of the outside of the test region besides showing typical propeller cavitation of the blade surface and tip-vortex types.

Free-stream type of propeller testing does not simulate properly the flow conditions encountered by a propeller. For ships, as well as submarines and torpedoes, propellers are always placed at the rear of the body and hence operate in a very disturbed flow region. However, the naval architect is forced to adopt this free-stream type of test due to the difficulty of simulating a ship's wake. For a proper test he would have to model a ship form behind which to test his propeller. To do this and keep the propeller size reasonable, say, 6 to 12 in., would result in a huge ship model; while to keep the size of the ship model reasonable would result in an impractically small propeller.

One approach to achieve the proper flow conditions for propeller operation has been made. This is in the Garfield Thomas Water Tunnel built by the Navy at The Pennsylvania State University for studies on torpedo propellers. Because of the prime interest in torpedoes, which are much smaller than ships or sub-

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

FIG. 1 PROPELLER CAVITATION TEST IN 12-IN. TUNNEL AT DAVID TAYLOR MODEL BASIN
(Official U. S. Navy Photo.)

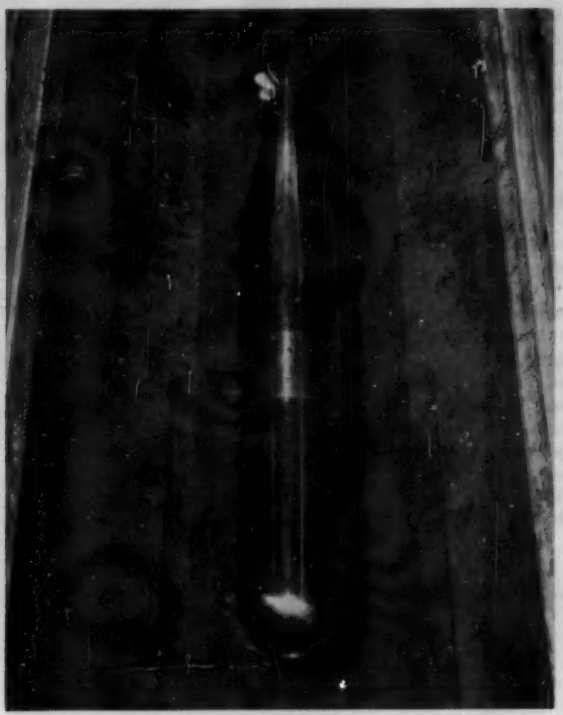


FIG. 2 MODEL TORPEDO AND PROPELLER INSTALLED IN WORKING SECTION OF 48-IN. WATER TUNNEL

FIG. 3 COUNTERROTATING PROPELLER BEING TESTED ON REAR OF MODEL TORPEDO

FIG. 4 WORKING AREA OF HIGH-SPEED WATER TUNNEL AT CALIFORNIA INSTITUTE OF TECHNOLOGY

marines, it has been possible to construct a tunnel large enough to include both the body and the propeller. Even though torpedoes are about 22 in. diam, the resultant tunnel is one of the largest in existence, being 100 ft long with a circular working section 4 ft in diameter.

The 8-in-diam torpedo model with its propeller is shown installed in the working section of this tunnel in Fig. 2. A typical cavitation picture of propellers tested on this model is shown in Fig. 3. This shows a pair of counterrotating propellers, producing blade-surface cavitation, tip-vortex cavitation, and a third type which had not been noted in free-stream cavitation tests, hub-vortex cavitation. This tunnel is used in many other types of tests besides propeller testing.

Nonpropeller Test Tunnels. Nonpropeller research and testing tunnels have been used for the study of the flow characteristics of hydrofoils, for the measurement of forces on bodies, and in the design of missile noses for hydroballistic applications.

The use of water tunnels for studies on bodies, and specifically noses of bodies, is best exemplified by water tunnels at the California Institute of Technology and the State University of Iowa. Two of these tunnels were built during the war and much of their war work was devoted to this type of study. At Cal Tech the 14-in. high-speed water tunnel is equipped with a three-compo-

nent balance for measuring forces on bodies such as torpedo models, hydrofoils, guide vanes, etc.

Fig. 4 shows the working-section region of this tunnel with the balance underneath it. The study of noses for hydroballistic bodies at the University of Iowa was concerned with cavitation conditions for a geometric series of model noses. Complete bodies were not studied and no force measurements were involved. Only cavitation conditions and pressures distributions were of interest.

In developing suitable nose shapes for hydroballistic devices such as torpedoes, another approach besides studying geometric shapes was developed following Reichardt's war work in Germany. In this method a blunt object is placed in the tunnel and the cavitation index lowered to the point where a clearly defined cavity is formed behind the body. The outline of this cavity is photographed and used to specify the contour of the forward portion of the body. The resultant shape should have a contour with essentially constant pressure distribution.

At the David Taylor Model Basin this technique has been applied by Eisenberg (D8) to determine nose contours of specified cavitation index, i.e., the index at which the photograph is taken. Fig. 5 shows time and instantaneous exposures of the cavity behind a disk.

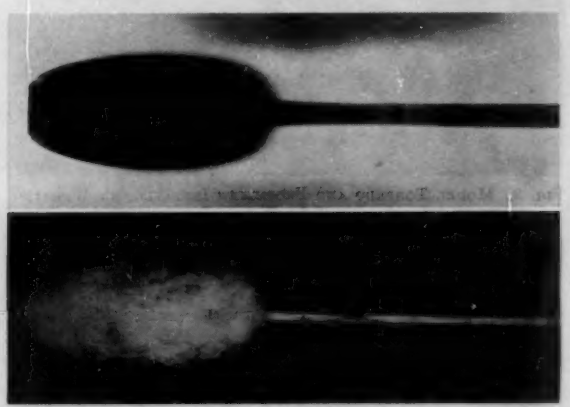


FIG. 5 CAVITATION BUBBLE FORMED BEHIND A DISK AT DAVID TAYLOR MODEL BASIN
(Official U. S. Navy Photo.)

Another technique, also introduced by Reichardt, involves blowing air into the space behind a disk or other shape. If this air pressure is considerably above the vapor pressure an effectively lower index is obtained (p_v is now replaced by p_{air} in the equation for σ) and a clear cavity is formed. Such a water-tunnel test is shown in progress in Fig. 6 which is a photograph of the working section of the Cal Tech free-surface water tunnel.

The use of water tunnels for flow studies on hydrofoils is probably the second oldest application of water tunnels. The work of Ackeret at Göttingen, Germany, published in 1930, describes work on hydrofoil cavitation in a research study initiated in 1925. Later work in this same tunnel is reported by Walchner on circular-arc airfoils. Some of his results showing the effect of cavitation on foil characteristics are given in Fig. 7. The initial improvement of section lift-drag ratio as cavitation becomes more severe (σ decreasing) helps explain the initial increase often noted in the efficiency of pumps or turbines as the amount of cavitation is increased.

Another series of hydrofoil studies in a water tunnel, illustrating the kind of work that can be done, were reported by Daily (D12), in 1949. This work was done in the Cal Tech 14-in.

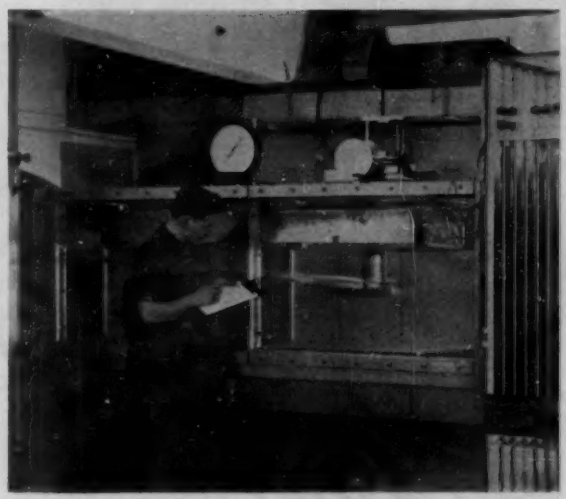


FIG. 6 WORKING AREA OF FREE-SURFACE WATER TUNNEL AT CALIFORNIA INSTITUTE OF TECHNOLOGY

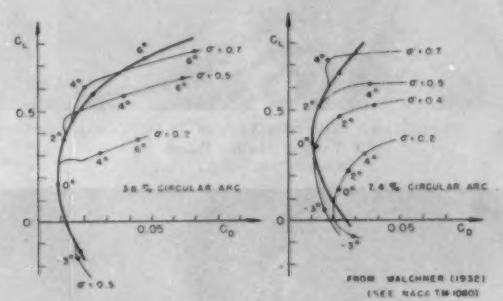


FIG. 7 POLAR DIAGRAMS FOR CIRCULAR-ARC HYDROFOILS UNDER CAVITATING CONDITIONS

water tunnel and included cavitation-inception measurements and force and moment measurements. Using the water-tunnel balance, Daily measured the lift and moment characteristics of the hydrofoil under noncavitating conditions. This could have been done with a variation in the cavitation index such as Walchner did, to obtain information on cavitation disturbance of the forces on the hydrofoil. Hydrofoil test work has been reported by Numachi who studied more conventional foil shapes than the circular arcs used by Walchner. The most recent work on this subject is that reported by Numachi in 1953 (B23) on hydrofoils in cascade as affected by the degree of cavitation.

Nonpropeller tunnels at the State University of Iowa also have been used to study the flow past boundary discontinuities such as offsets or slots. Slots are of interest as they tend to cause cavitation at the entrance of conduits or other hydraulic structures (A13).

A final research use for water tunnels is in the study of cavitation fundamentals. Here interest lies in trying to learn more about cavitation, i.e., to study what factors affect its inception and how it behaves afterward. Important studies along this line have been made in the 14-in. tunnel at Cal Tech. The high-speed motion-picture study of the growth and collapse of a bubble on the nose of a body was presented and discussed by Knapp and Hollander in 1948 (D10), and has been analyzed further by Plesset (D13).

More recent studies at Cal Tech have centered around the effect of body size or scale on the cavitation-inception condition for simple nose shapes. These studies have been reported by Parkin (D18) and have been presented also in a paper by Knapp. A variation in the index, both with tunnel speed and model size, has been noted. Such studies are still continuing and were supplemented recently by similar work in the Ordnance Research Laboratory's 48-in. water tunnel at The Pennsylvania State University. The latest word on this subject is presented in a 1954 paper by Kermeen, McGraw, and Parkin (D21).

Experimental Water Tunnels. Experimental or model tunnels are more rare than the research type but there have been a half dozen of these in the country and they have yielded information of considerable value. Although experimental water tunnels are constructed to obtain information for tunnel designers, they often produce information on flow in hydraulic components of value to designers of pipe lines and conduits. Such tunnels were built by the NACA in 1928, The Pennsylvania State College in 1946 (ORL), the University of Minnesota (St. Anthony Falls Hydraulic Lab) starting in 1947, and the Massachusetts Institute of Technology (Hydrodynamics Lab). Further discussion of these is presented in the referenced report.[2]

Description of Water Tunnels

Types of Water Tunnels

As in the case of air tunnels, water tunnels may be typed according to several criteria. These involve the nature of the flow circuit, the nature of the operating cycle, and the design of the test region. According to the first criterion one can have either a closed or open circuit for the tunnel. Most of the water tunnels have closed circuits because they do not have an infinite sink to dump the fluid into and to draw it out of as one may have for an air tunnel. Another distinction between types of water tunnels has to do with their operating cycle. Most water tunnels are run continuously. This means that a pump in the circuit is used to circulate the water as long as the tunnel is in operation. Noncontinuous or blowdown tunnels are becoming more common for air tunnels but are rare in water tunnels. The 10-in. free-jet

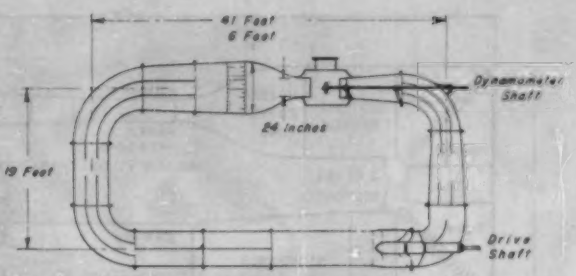


Fig. 8 24-In. Water Tunnel at David Taylor Model Basin

water tunnel at the University of Minnesota is a gravity-flow tunnel which probably also can be classed as a noncontinuous-flow tunnel although the high-pressure supply is a river in this case.

As most water tunnels are closed-circuit, continuous-flow tunnels, the main distinction between them has to do with the nature of the jet in the test region. Just as for air tunnels, they can be either closed or open-jet working sections. In the closed-jet working section the stream of fluid in which the tests are to be performed is bounded by solid walls while in the open-jet working section the fluid is surrounded by a body of relatively still fluid. Both types of working sections are used as noted in Table 1. The reason for choosing one or the other type of working section is essentially the same as in air-tunnel construction.

In a closed-jet working section a sizable wall-interference effect occurs when bodies are tested in it. This effect is reduced greatly when an open jet is used; however, the usual open-jet design is subject to the restriction that it cannot be made much more than one diameter long before instability problems occur.

A third type of working section of a water tunnel is the free-jet one in which the stream of water flows through air under relatively low pressure. This type of tunnel was first used by Reichardt in Germany during the war to obtain extremely low cavitation indexes. It is best suited for studies under extreme cavitation conditions (C6).

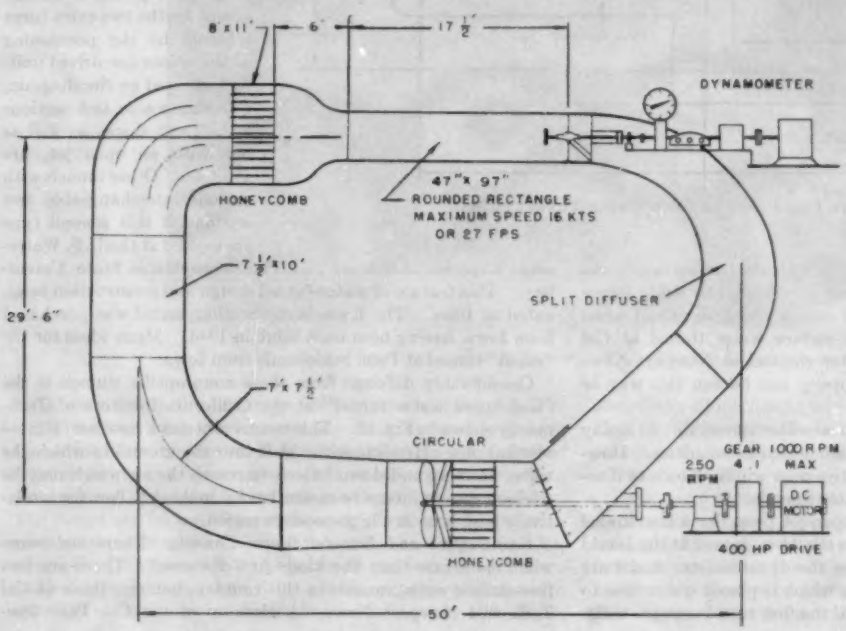


Fig. 9 No. 2 Hamburg Tunnel—1940 (Now at Haslar, England)

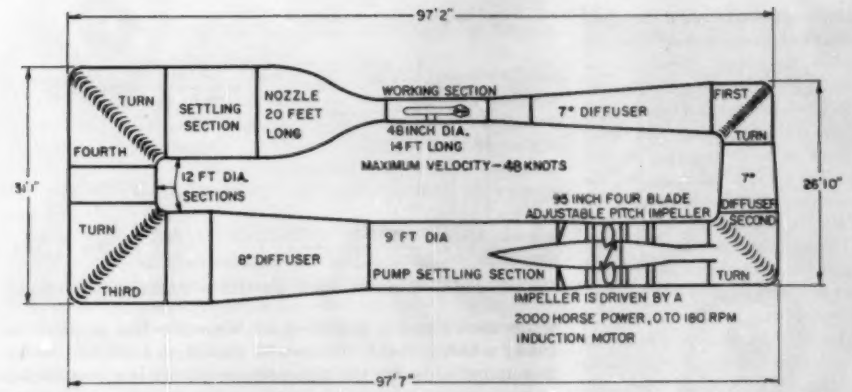


FIG. 10 ORDNANCE RESEARCH LABORATORY WATER TUNNEL AT THE PENNSYLVANIA STATE UNIVERSITY

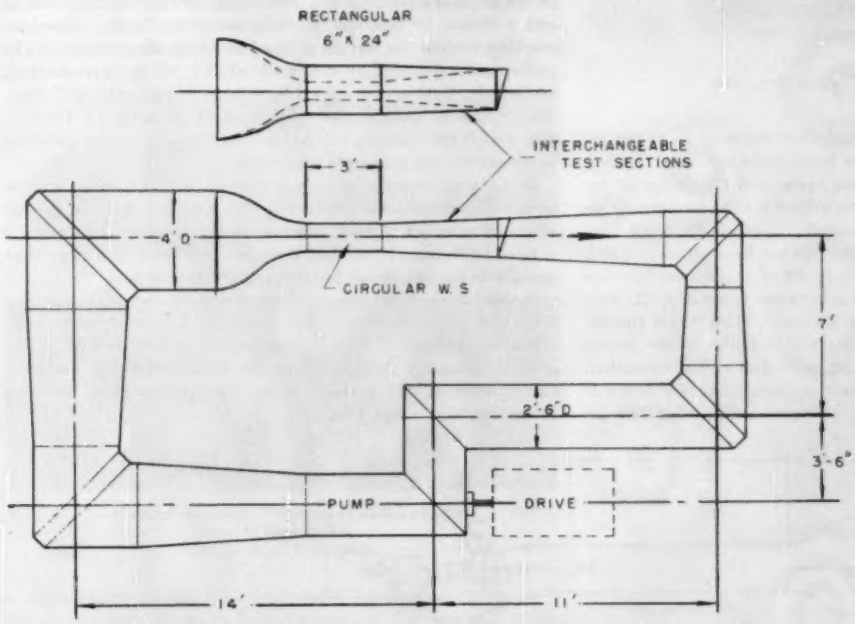


FIG. 11 STATE UNIVERSITY OF IOWA WATER TUNNEL—1949

A final type of water tunnel is that with the free-surface working section. In this type of tunnel it is desired to study forces acting on a body when it is on or near a free surface and when cavitation is involved. The free-surface water tunnel at Cal Tech is of this type and the water channel at Newport News Shipbuilding and Dry Dock Company can be run this way or with a closed working section.

There is no such thing as a typical water tunnel for, as in any research facility, each one has its own design peculiarities. However, some of the more representative ones will be shown to illustrate the important features of water tunnels.

Propeller Tunnels. A typical open-jet propeller water tunnel is illustrated in Fig. 8 which shows the 24-in. tunnel at the David Taylor Model Basin. In this case the dynamometer shafts are brought out through the first turn which is placed quite close to the test region. This placement of the first turn is not too satisfactory. It is, however, a common feature of propeller tunnels. A closed jet-propeller tunnel of the Hamburg Model Basin (HSVA) is shown in Fig. 9. This tunnel was built to permit testing ship models with propellers and therefore has an extremely large test section. Although it is indicated as being built in 1940, it does not seem to have operated yet due to difficulties resulting from the war.

In the case of the 48-in. water tunnel at The Pennsylvania State University, a powered torpedo model permitted the use of more rational diffuser and circuit design as it was not necessary to bring out any external shafting. The circuit diagram for this tunnel, Fig. 10, shows the design details and principal dimensions of this large tunnel. It is not specifically a propeller tunnel. However, it can test propellers in the free-stream condition as large as can be tested in any other tunnel.

Nonpropeller Water Tunnels. These usually have circuit diagrams similar to that just shown. Thus the circuit for the 1949 water tunnel at the State University of Iowa, Fig. 11, is similar except for the two extra turns required by the positioning of the motor (or drive) unit. As indicated on the diagram, interchangeable test sections of different shape, as well as one with an open jet, are available. Other tunnels with various interchangeable test sections of this general type are located at the U. S. Waterways Experiment Station and The Pennsylvania State University. This feature of water-tunnel design and construction originated at Iowa. The Experiment Station tunnel was purchased from Iowa, having been built there in 1944. Many ideas for the "small" tunnel at Penn State came from Iowa.

Considerably different from these nonpropeller tunnels is the "high-speed water tunnel" at the California Institute of Technology shown in Fig. 12. This tunnel contains a resorber (B18)—vertical pipe extending some 58 ft into the ground in which the water flows up and down twice—to resorb the air which may be released during intensive cavitation by making it flow for a relatively long time in a high-pressure region.

Free-Surface and Free-Jet Water Tunnels. These are somewhat more rare than the kinds just discussed. There are two free-surface water tunnels in this country, namely, those at Cal Tech and Newport News. A close-up of the Cal Tech free-

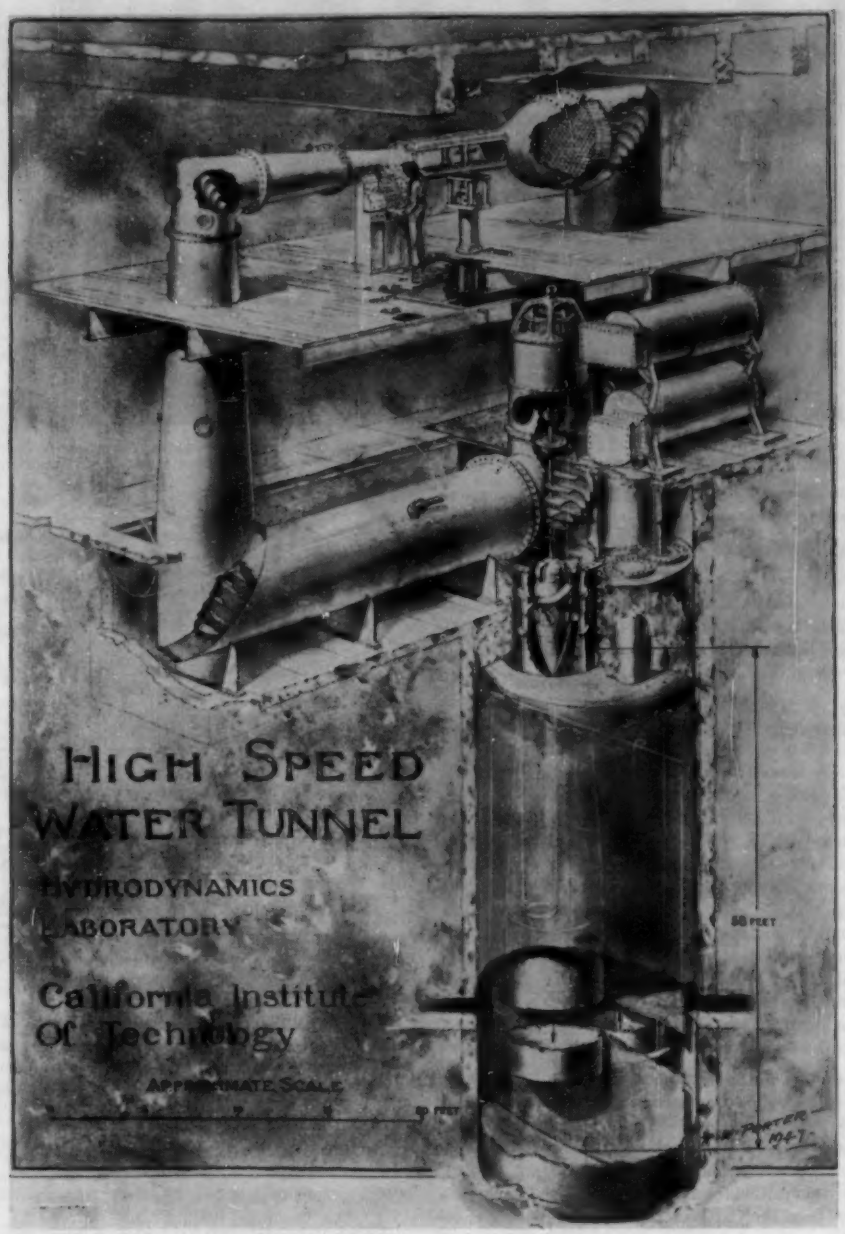


FIG. 12 14-IN. HIGH-SPEED WATER TUNNEL AT CALIFORNIA INSTITUTE OF TECHNOLOGY

surface water-tunnel working section was shown in Fig. 6. Owing to the need for a special air or gas separator used to remove the air introduced in the working section during cavity and other tests requiring gas injection, the diffuser following the working section bears little resemblance to any other water-tunnel diffuser. The tunnel can be run either with atmospheric pressure or with variable pressure over the free surface. In the design of such a tunnel the problems of wave formation in the working section are naturally very important (B19). A photograph of the water tunnel or water channel at the Hydraulics Laboratory of the Newport News Shipbuilding and Dry Dock Company is shown in Fig. 13. This channel can be run either closed or with a free surface and is a quarter-scale model of a larger facility which it is hoped to build.

Even rarer than free-surface tunnels are the free-jet tunnels for tests at extremely low cavitation numbers. The 10-in. tunnel at the University of Minnesota, St. Anthony Falls Hydraulic Laboratory (B24) is the only operating one in this country. It has a jet length of 40 in. partially restrained by viewing windows. It is capable of reaching cavitation indexes as low as 0.01. Fig.

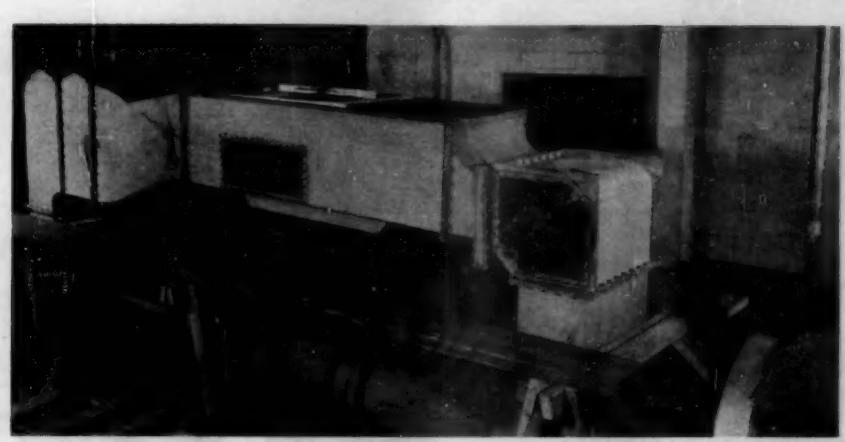


FIG. 13 WATER CHANNEL OF NEWPORT NEWS SHIPBUILDING AND DRY DOCK COMPANY

14 illustrates the manner in which the river flow is passed through this tunnel and the flow conditions controlled.

HYDRAULIC USES FOR WATER TUNNELS

In discussing the purposes for which water tunnels have been built and the various kinds and types of water tunnels, the hydraulic uses to which they may be applied have been noted. In summary, a listing is presented of the several nonpropeller testing uses of water tunnels:

1 Measurement of hydrofoil characteristics—determination of cavitation-inception conditions, and lift and drag under cavitating conditions.

2 Studies of nose shapes—study of the cavitation-inception conditions for nose profiles, and the development of suitable contours for prescribed cavitation indexes.

3 Studies of aerodynamic shapes at high Reynolds numbers (without cavitation or compressibility effects).

4 Studies of aerodynamic shapes for near-sonic use by the cavitation analogy.

5 Studies of hydraulic structures components—study of the cavitation conditions near slots, and the flow conditions around baffle piers and piles.

6 Studies of flow in components of hydraulic systems and general studies (as in experimental tunnels).

7 Studies of cavitation fundamentals.

These uses for water tunnels have been discussed to some extent. The study of hydrofoil characteristics is of course of great value to hydraulic machinery designers as well as to propeller designers. The use of a water tunnel for hydraulic structures tests is increasing. Certainly there has been a great, and needed, increase in studies of cavitation fundamentals. Typical researches along this line have been the Cal Tech high-speed movies of cavitation bubbles, the studies of pressure distribution, and cavitation inception on noses at the University of Iowa, and the cavitation-inception scaling studies at Cal Tech.

ACKNOWLEDGMENT

The illustrations used in this paper have come from many sources. Special acknowledgment is due to the David Taylor Model Basin for Figs. 1, 5, and 8; to the Hydrodynamics Laboratory of the California Institute of Technology for Figs. 4, 6, and 12; to the Iowa Institute of Hydraulic Research for Fig. 11; to the Hydraulics Laboratory of the Newport News Shipbuilding and Dry Dock Company for Fig. 13; to the St. Anthony Falls Hydraulic Laboratory of the University of Minnesota for Fig. 14;

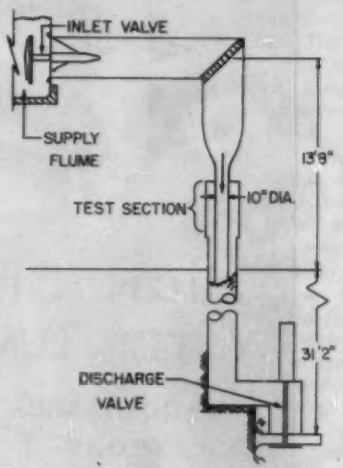


FIG. 14 LAYOUT OF FREE-JET TUNNEL AT UNIVERSITY OF MINNESOTA

and to the Ordnance Research Laboratory of The Pennsylvania State University for Figs. 2, 3, and 10. The compilation of the material summarized in this paper was supported by the U. S. Navy Bureau of Ordnance under Contract NOrd 7958 with The Pennsylvania State University.

A SHORT BIBLIOGRAPHY ON CAVITATION AND WATER TUNNELS

A, Cavitation and Associated Hydraulic Problems

1 "An Annotated Bibliography of Cavitation," by F. A. Raven and A. M. Feiler, David Taylor Model Basin, Report R-81, December, 1947.

2 "The Causes of the Racing of Screw Streamers, Investigated Theoretically and by Experiment," by O. Reynolds, Transactions of the Institution of Naval Architects (INA), 1873; Scientific Papers I, pp. 51–58.

3 "On the Effect of Immersion on Screw Propellers," by O. Reynolds, Trans. INA, 1874; Scientific Papers I, pp. 78–80.

4 "Experiments Showing the Boiling of Water in an Open Tube at Ordinary Temperatures," by O. Reynolds, British Association for the Advancement of Science, Report, p. 564, 1894; Scientific Papers II, pp. 558–587.

5 "On the Formation of Cavities in Water by Screw Propellers at High Speeds," by S. W. Barnaby, Trans. INA, vol. 38, 1897, pp. 139–144.

6 "The Application of the Compound Steam Turbine to the

Purpose of Marine Propulsion," by C. A. Parsons, Trans. INA, vol. 38, 1897, pp. 232–242.

6a "Experimental Apparatus Showing Cavitation in Screw Propellers," by C. A. Parsons, Transactions of the North East Coast Institution of Engineers and Shipbuilders, vol. 29, 1912–1913, pp. 300–302.

7 "Investigations Into the Causes of Corrosion or Erosion of Propellers," by C. A. Parsons and S. S. Cook, *Engineering* (London), April 18, 1919, pp. 515–519.

8 "The Turbine Testing Laboratories at Verkstaden, Kristinehamn, Sweden," by J. R. Freeman, Hydraulic Laboratory Practice, ASME, 1929, pp. 631–634.

9 "Experimental and Theoretical Investigation of Cavitation in Water," by J. Ackeret, Report of the Kaiser Wilhelm Institut für Stromungsforschung, Göttingen, Germany, January, 1930; *Technische Mechanik und Thermodynamik*, vol. 1, pp. 1–22; David Taylor Model Basin, Translation 20, February, 1936; NACA, TM 1078, July, 1945.

10 "Cavitation Testing of Model Hydraulic Turbines and Its Bearing on Design and Operation," by L. M. Davis, Trans. ASME, vol. 57, 1935, pp. 455–462.

11 "Erosion and Cavitational Erosion," by P. de Haller, Handbuch der Werkstoffprufung, Berlin, Germany, vol. 2, 1939, pp. 471–488; David Taylor Model Basin, Translation 188, April, 1951.

12 "Model and Prototype Tests" in "Trend in Hydraulic Turbine Practice—A Symposium," by L. M. Davis, Trans. ASCE, vol. 106, 1941, pp. 353–367.

13 "Cavitation in Hydraulic Structures—A Symposium," Trans. ASCE, vol. 112, 1947, pp. 1–124.

14 "On the Mechanism and Prevention of Cavitation," by P. Eisenberg, David Taylor Model Basin, Report No. 712, July, 1950.

15 "Cavitation Mechanics in the Design of Hydraulic Equipment," by R. T. Knapp (James Clayton Lecture of The Institution of Mechanical Engineers—1952), *The Engineer* (London), April 25, 1952, pp. 571–574; May 2, 1952, pp. 592–596; May 9, 1952, pp. 627–628.

B, Water-Tunnel Descriptions

1 A Bibliography on Existing Water Tunnels (with chart and photostats of circuits), by J. F. Ripken, unpublished work at David Taylor Model Basin, May, 1945.

2 See reference (A6).

3 See reference (A9).

4 "Experiments With a Model Water Tunnel," by E. N. Jacobs and I. H. Abbott, NACA, TN 358, December, 1930.

5 "The Variable Pressure Water Tunnel of the U. S. Experimental Model Basin," by H. E. Saunders, Trans. of The Society of Naval Architects and Marine Engineers (SNAME), vol. 38, 1930, pp. 205–217.

6 "The Cavitation Tank of the Hamburgische Schiffbau-Versuchsanstalt," by H. Lerbs, Werft-Reederei-Hafen, vol. 12, No. 11, June, 1931; David Taylor Model Basin, Translation 14, February, 1936.

7 "New Cavitation Research," by H. Lerbs, Werft-Reederei-Hafen, vol. 20, No. 16, August, 1936; David Taylor Model Basin, Translation 89, January, 1940.

8 "Propeller Testing Tunnel at the Massachusetts Institute of Technology," by F. M. Lewis, Trans. SNAME, vol. 47, 1939.

9 "Variable Pressure Water Tunnel—Netherlands Model Testing Basin," by J. G. Koning, from Report 47 of the Netherlands Model Basin, Marineblad, June, 1941; David Taylor Model Basin, Translation 98, December, 1941.

10 "The Variable Pressure Water Tunnels at the David Taylor Model Basin," by A. G. Mumma, Trans. SNAME, vol. 49, 1941.

11 "Kraftmessungen an vier Flugelprofilen bei Kavitation," by F. Numachi, Technology Reports of the Tôhoku University, Japan, Vol. XIII, No. 3, 1943; also *VDI Forschung*, vol. 11, 1940, p. 303.

12 "The Test Tunnel for Ship Propellers at Karlstads Mekaniska Werkstad, Karlstad, Sweden," by H. Edstrand, Kristinehamn, Sweden, 1945; translation of part of an article in *Teknisk Tidskrift*, vol. 20, 1944.

13 "Influence of Reynolds Number on the Separation (Cavitation) Flow," by V. A. Konstantinov, Izvestiia Akademii Nauk SSSR, Otdelenie Teckhnickeskikh Nauk, No. 10, 1946, pp. 1355–1373; David Taylor Model Basin, Translation 233, November, 1950.

14 "On Cavitation Tunnels for Small Cavitation Numbers," by H. Reichardt, Kaiser Wilhelm Institut für Stromungsforschung, February 12, 1945; Translation by Royal Aircraft Establishment, London, England, November, 1945.

15 "The Water Tunnel as a Tool in Hydraulic Research," by J. W. Daily, Proceedings of the Third Hydraulics Conference, 1946; State University of Iowa, Studies in Engineering, Bulletin 31, 1947, pp. 179–191.

16 "Pressure Distribution and Cavitation on Submerged Bodies," by J. S. McNown, ibid., pp. 198–199.

17 "Hydrodynamic Design of the 48-Inch Water Tunnel at The Pennsylvania State College," by D. Ross, J. M. Robertson, and R. B. Power, Trans. SNAME, vol. 56, 1948, pp. 5–29.

18 "The Hydrodynamics Laboratory of the California Institute of Technology," by R. T. Knapp, J. Levy, J. P. O'Neill, and F. B. Brown, Trans. ASME, vol. 70, 1948, pp. 437–457.

19 "The Hydrodynamics of the Free Surface Water Tunnel," by J. P. O'Neill, Hydrodynamics Laboratory, California Institute of Technology, Report N-65, May, 1949.

20 "Garfield Thomas Water Tunnel Operations," by R. B. Power, J. M. Robertson, and D. Ross, et al., Ordnance Research Laboratory, School of Engineering, The Pennsylvania State University, Report NOrd 7958-211, May, 1951.

21 "Piping Big Water Tunnel for Naval Research," by J. M. Robertson and R. L. Steele, *Heating, Piping & Air Conditioning*, November, 1951, pp. 86–90.

22 "New Hydrodynamic Research Facilities at the Ordnance Research Laboratory," by A. F. Lehman, paper presented at the Ninth Underwater Ballistics Conference, State College, Pennsylvania, July, 1952.

23 "Cavitation Tests on Hydrofoils in Cascade," by F. Numachi, Trans. ASME, vol. 75, 1953, p. 1257.

24 "Description of a Ten-Inch Free-Jet Water Tunnel," by C. D. Christopherson, St. Anthony Falls Hydraulic Laboratory, University of Minnesota, Project Report No. 35, January, 1953.

C, Research on Tunnels and Components

1 "The Experimental Water Tunnel at The Pennsylvania State College," by J. M. Robertson, D. Ross, and J. R. Villemonte, Ordnance Research Laboratory, School of Engineering, The Pennsylvania State University, Report NOrd 7958-89, April, 1948.

2 "Design of Vaned Turns for a Large Water Tunnel," by A. J. Turchetti and J. M. Robertson, paper presented at ASME meeting, June, 1948.

3 "Air Resorption in Water Tunnels," by F. B. Brown, Hydrodynamics Laboratory, California Institute of Technology, Report N-62, March, 1949.

4 "Design Studies for a Closed Jet Water Tunnel," by J. F. Ripken, St. Anthony Falls Hydraulic Laboratory, University of Minnesota, Technical paper No. 9, Series B, August, 1951.

5 "The Unsteady-Flow Water Tunnel at the Massachusetts Institute of Technology Hydrodynamics Laboratory, Department of Civil and Sanitary Engineering, Massachusetts Institute of Technology, Technical report No. 2, May, 1951.

6 "Experimental Design Studies on Free-Jet Water Tunnels," by C. D. Christopherson, St. Anthony Falls Hydraulic Laboratory, University of Minnesota, Report ONR No. 25, September, 1951.

7 "Effect of Entrance Conditions on Diffuser Flow," by J. M. Robertson, D. Ross, Proceedings of the ASCE, Separate No. 141, July, 1952.

8 "Water-Tunnel Transition—Section Flow Studies," by D. Ross, Ordnance Research Laboratory Report NOrd-7958-285, June, 1954.

9 "Turbulent Flow at the Entrance Region of a Pipe," by D. Ross. Paper No. 54—A-89.

D, Hydraulic Research for Water Tunnels

1 See reference (A9).

2 "Profile Measurements During Cavitation," by O. Walchner. Paper presented at Conference on Hydromechanische Probleme des Schiffsantriebs, Hamburg, Germany, 1932, David Taylor Model Basin, Translation 21, 1936; NACA, TM 1060, 1944.

3 "Profilmessungen bei Kavitation: 1—Mitteilung Kraftmessungen an vier Flugelprofilen bei Kavitation. 2—Mitteilung Flugelprofilen bei Ruckwartiger Anstromung. 3—Mitteilung Kraftmessungen an den Spaltflugelprofilen bei Kavitation," by F. Numachi, Technology Reports of the Tôhoku Imperial University, Japan, Vol. XIII, No. 3, 1943, pp. 409–524.

4 "Ringhauben mit Kavitationskontur," by H. Reichardt, Kaiser Wilhelm Institut für Stromungsforschung, December, 1944.

5 "On the Determination of Favorable Shapes of Bodies for High Speed by Cavitation Tests," by H. Reichardt, Joint Intelligence Objectives Agency (British Interrogation of German Scientists, No. 17, CGD-495), June 26, 1946.

6 See reference (B13).

7 See reference (B15).

8 "A Cavitation Method for the Development of Forms Having Specified Critical Cavitation Numbers," by P. Eisenberg, David Taylor Model Basin, Report 647, September, 1947.

9 "Water Tunnel Investigations of Steady-State Cavities," by

P. Eisenberg and H. L. Pond, ibid., Report 668, October, 1948.
10 "Laboratory Investigations of Mechanism of Cavitation," by R. T. Knapp and A. Hollander, Trans. ASME, vol. 70, 1948, p. 419.
11 "Cavitation and Pressure Distribution—Head Forms at Zero Angle of Yaw," by H. Rouse and J. S. McNown, State University of Iowa, Studies in Engineering, Bulletin 32, 1948.
12 "Cavitation Characteristics and Infinite Aspect-Ratio Characteristics of a Hydrofoil Section," by J. W. Daily, Trans. ASME, vol. 71, 1949, p. 269.
13 "The Dynamics of Cavitation Bubbles," by M. S. Plesset, *Journal of Applied Mechanics*, Trans. ASME, vol. 71, 1949, p. 277.
14 "Cavitation Tests on Six Profiles for Blade Elements," (in Japanese), by F. Numachi, K. Tsunoda, and I. Chida, Reports of the Institute of High Speed Mechanics, Tôhoku University, Sendai, Japan, vol. I, No. 1, March, 1949.
15 "Cavitation Tests on Clark Y Profiles of Several Thickness Ratio" (in Japanese), by F. Numachi, K. Tsunoda, and I. Chida, ibid., No. 2.
16 "Cavitation Tests on Special Profiles for Blade Elements (X-Profiles)," (in Japanese), ibid., No. 3. (14, 15, and 16 are translated in TM 19.4210, Ordnance Research Laboratory, The Pennsylvania State University, July, 1954.)
17 See reference (C5).
18 "Scale Effects in Cavitating Flow—A Preliminary Report," by B. R. Parkin, Hydrodynamics Laboratory, California Institute of Technology, Report No. 21-7, December 28, 1951.
19 See reference (B23).
20 "Aerodynamics of Propulsion," by D. Kucheman and J. Weber, McGraw-Hill Book Company, Inc., New York, N. Y., 1953, pp. 103–105, 241–242, 274.
21 "Mechanism of Cavitation Inception and the Related Scale-Effects Problem," by R. W. Kermeen, J. T. McGraw, and B. R. Parkin, Trans. ASME, vol. 77, 1955, pp. 533–541.
22 "Water Tunnel—Tool in Aerodynamic Research," by J. M. Robertson, *Aero Digest*, vol. 69, November, 1954.

Discussion

J. F. Ripken.[4] The demand of man for an ever-increasing speed of motion through fluids requires that the tools of the engineer keep pace with his problems. The author is to be congratulated on his very useful exposé of the water tunnel as an engineering tool for the examination of cavitation resulting from high-speed motions in liquids.

With regard to the mentioned free-jet tunnel at the St. Anthony Falls Hydraulic Laboratory, it may be worthy of note that this tunnel has recently produced test data at sigma values some 20 per cent ($\sigma = 0.008$) lower than those cited by the author, and there are indications that a free-jet tunnel can be built to produce values even lower than this. This tunnel is of an open-circuit, continuous-flow type.

It is also noteworthy that an additional new form of test section of European origin has been tested recently in the 6-in. water tunnel at this laboratory. This so-called slotted-wall test section consists essentially of an open-jet chamber fitted with an inner closed-jet sleeve having numerous longitudinal slots around its periphery. The restrained communication between the jet fluid and the chambered fluid of this hybrid test section produces a hybrid performance characterized by or permitting small wall-interference effects, an essentially flat longitudinal pressure gradient, a substantial test-section length, and a high stream velocity.

It may be well to point out to the nonspecialist that the water tunnels described by the author are primarily for the study of submerged moving bodies. Special modifications of these variable-pressure, recirculating, cavitation-test facilities have, however, long been used as test stands by the makers of pumps and turbines. Special modifications also have been built to house model hydraulic structures which are stationary geometric bodies (spillways, energy dissipators, etc.) exposed to high-velocity flow in their prototype form.

[4] Professor of Hydromechanics, St. Anthony Falls Hydraulic Laboratory, University of Minnesota, Minneapolis, Minn.

Author's Closure

Research facilities are inherently individualistic and the use to which they are put often leads to modification in their nature. Consequently, it is difficult to place them in firm categories. This difficulty is apparent in the portion of the paper concerned with description of "typical" water tunnels. In fact, before one type, propeller tunnel, has been covered it is noted that the largest example is typical of nonpropeller tunnels.

In his discussion, Mr. Ripken calls attention to the similarity of water tunnels to the test stands used by pump and turbine manufacturers. In the report[2] which formed the basis of this paper, reference was made to an early (1924) turbine test facility (A8) built in Sweden. It was noted that this apparatus might not be a water tunnel. As of today, the narrow zone of distinction between the two types of facility has disappeared. In oral discussion of the paper, mention was made (by S. Logan Kerr) of a Canadian tunnel designed to test pumps or turbines. Although not cognizant of this particular facility, the author has been aware of a tunnel being built in Great Britain in which the working section can be removed and a turbine installed in its place for cavitation testing. Thus in the same facility we have a turbine test stand and a water tunnel.

The slotted-wall test section mentioned by Mr. Ripken adds greatly to tunnel flexibility. Such a section has been installed in one of the water tunnels at the University of Iowa. It is interesting to note that although the form this takes in a water tunnel is unique, the concept was adapted by the British from a supersonic wind-tunnel application.

The author appreciates greatly the discussion by Mr. Ripken. The bibliography which he prepared in 1945 (B1) not only helped introduce the author to water tunnels, but was invaluable in the preparation of this paper.

An Extension of the Theory of Water Hammer

By R. SKALAK,[1] NEW YORK, N. Y.

The usual theory of water hammer which is used to compute the magnitude and velocity of pressure waves in a pipe filled with an elastic fluid was given by Joukowsky in 1898. This theory predicts that pressure waves travel without change of shape. In the present paper, water-hammer waves are considered without some of the simplifications made in the Joukowsky theory. The equations of motion for a thin cylindrical tube are used. Only the case where the velocity of sound in the material of the pipe wall is greater than in the fluid will be considered. Much of the analysis would be valid in the opposite case, but the relative speeds of the various waves would differ.

Nomenclature

The following nomenclature is used in the paper:

c = velocity of sound in fluid
c_o = Joukowsky water-hammer velocity (see Equation [1])
K = bulk modulus of fluid
a = radius of tube
h = thickness of tube wall
E = Young's modulus of tube-wall material
p = pressure in fluid
v = axial velocity of fluid
x = distance axially along tube
t = time
u = axial displacement of tube
w = radial deflection of tube wall (+ outward)
ρ_0 = initial density of fluid
r = radial distance
ϕ = velocity potential
m = mass of tube per unit surface area
ν = Poisson's ratio for tube wall
$c_0 = \sqrt{\frac{Eh}{m(1-\nu^2)}}$ = velocity of sound in tube wall
q = force per unit area on tube wall
ξ, Ω = transform variables
n, s = integers
N = see Equation [23]
Δ = see Equation [24]
J_0, J_1 = Bessel functions
$i = \sqrt{-1}$
D = see Equation [30]
F = see Equation [36]
G = see Equation [37]
B = see Equation [38]
$c', \bar{c}$ = arbitrary constant velocities
c_1, c_2 = phase velocities
z' = see Equation [48]
η = variable of integration
β = see Equation [50]
I, H = integrals (see Equation [53])
L = length of wave front
$R = c_0^2/c^2$
$A = \rho_0 a/m$

[1] Assistant Professor of Civil Engineering, Columbia University.
Contributed by the Hydraulic Division and presented at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–22, 1955, of The American Society of Mechanical Engineers.
Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 19, 1954. Paper No. 55—S-18.

Introduction

The usual theory of water hammer which is used to compute the magnitude and velocity of pressure waves in a pipe filled with an elastic fluid was given by Joukowsky (8)[2] in 1898. This theory predicts that pressure waves travel without change of shape at a velocity c_o

$$c_o = \frac{c}{\sqrt{1 + \frac{2Ka}{Eh}}} \quad \ldots\ldots [1]$$

In the derivation of this formula it is assumed that the pressure is uniform across any section of the pipe and that the deflection of the pipe is equal to the static deflection due to the instantaneous pressure in the fluid. The first assumption neglects, in effect, the inertial forces associated with radial motion of the fluid. The second neglects the mass of the pipe wall and the longitudinal and bending stresses in the pipe wall.

For the case of an infinite train of sinusoidal pressure waves in a tube filled with fluid, much work has been done without resorting to the above assumptions. In 1878 Korteweg (9) showed that the phase velocity of sinusoidal waves varies with the wave length when the radial inertia of the fluid and the mass of the pipe wall are taken into account. Korteweg also gave the expression for c_o, Equation [1], as an approximation of the phase velocity for long wave lengths. The work of Korteweg was extended in 1898, by Lamb (10) who considered the pipe as an elastic membrane thus including the effect of longitudinal stresses in the pipe wall. Lamb showed there are two finite phase velocities as the wave length approaches infinity. One of these is very close to Korteweg's result c_o and the other is very nearly the velocity of sound in the material of the tube wall. Recent papers by Jacobi (6) and Thomson (14) include the flexural stiffness of the pipe wall. The inclusion of the bending stresses affects the phase velocities appreciably only for very short wave lengths.

In the present paper water-hammer waves are considered without some of the simplifications made in the Joukowsky theory. The equations of motion for a thin cylindrical tube are used, according to Flügge (11), with one additional term for the effect of rotatory inertia. The fluid is assumed to be elastic and inviscid but no assumption is made regarding its motion except that it is small.

Only the case where the velocity of sound in the material of the pipe wall is greater than in the fluid will be considered. Much of the analysis would be valid in the opposite case, but the relative speeds of the various waves would differ.

A Formal Solution by Integral Transforms

In this section a formal solution of a typical water-hammer problem is obtained in the form of a double integral. The solution

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

is reduced to terms of single real integrals whose evaluation is discussed in the next section.

The problem considered is to find the motion of a thin cylindrical tube filled with fluid starting from given initial conditions. The conditions, shown in Fig. 1, are that for $t = 0$

$$\left.\begin{array}{ll} p = 0, & v = 0, \text{ for } z > 0 \\ p = p_0, & v = v_0, \text{ for } z < 0 \\ u = 0, & w = 0, \text{ for all } z \end{array}\right\} \quad \ldots\ldots\ldots\ldots [2]$$

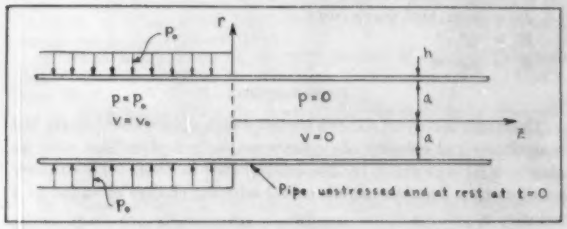


FIG. 1 INITIAL CONDITIONS

The velocity v_0 is assumed to be related to the pressure p_0 by

$$v_0 = \frac{p_0}{\rho_0 c} \quad \ldots\ldots\ldots\ldots [3]$$

This is the velocity that would exist behind a plane shock wave of pressure p_0 in an infinite body of fluid. Thus the initial conditions for the fluid correspond to a step shock wave moving in the positive direction. An external pressure equal to p_0 is assumed acting on the outside of the tube for $z < 0$ to prevent any motion of the tube at $z = -\infty$. This external pressure is assumed to remain on the tube for $z < 0$ for all t.

These initial conditions are intended to simulate the situation arising when a pipe leading from a reservoir is filled with fluid at rest below reservoir pressure and a valve at the reservoir end is suddenly opened. The part of the tube left of the origin in Fig. 1 is a convenient device for supplying fluid and energy in place of a reservoir. It is expected that the particular means of generating a water-hammer wave will not essentially affect its characteristics when the wave has progressed far enough down the pipe.

The motion of the fluid will be described by a velocity potential which must satisfy the equation

$$\nabla^2\phi = \phi_{rr} + \frac{1}{r}\phi_r + \phi_{zz} = \frac{1}{c^2}\ddot{\phi} \quad \ldots\ldots\ldots\ldots [4]$$

where the subscripts r and z denote derivatives.

The motion of the tube due to a radial force per unit area, $q(z, t)$, is governed by two equations

$$\frac{\ddot{w}}{c_0^2} - \frac{h^2}{12c_0^2}\ddot{w}_{zz} + \left(1 + \frac{h^2}{12a^2}\right)\frac{w}{a^2} + \frac{\nu}{a}u_z + \frac{h^2}{12}w_{zzzz} - \frac{h^2}{12a}u_{zzz} = \frac{q}{mc_0^2} \quad \ldots [5]$$

$$\frac{\ddot{u}}{c_0^2} - u_{zz} - \frac{\nu}{a}w_z + \frac{h^2}{12a}w_{zzz} = 0 \quad \ldots\ldots\ldots [6]$$

A detailed derivation of these equations is given in (12). They agree with the equations given by Flügge (11), except that the rotatory inertia term $\ddot{w}_{zz}$ has been added above.

The fact that the fluid and tube are expected to remain in contact is expressed by

$$\dot{w} = -\left[\frac{\partial\phi}{\partial r}\right]_{r=a} \quad \ldots\ldots\ldots\ldots [7]$$

The pressure exerted on the tube wall by the fluid is

$$[p]_{r=a} = \rho_0\left[\frac{\partial\phi}{\partial t}\right]_{r=a} \quad \ldots\ldots\ldots\ldots [8]$$

This is the value of q to be used for $z > 0$ in Equation [5]. For $z < 0$, the external pressure $-p_0$, must be added.

It is convenient to consider the solution of the problem as the sum of two parts.

The first part is the shock wave that would be produced by the initial conditions if the tube were rigid. The tube displacements w_1, u_1, for this part are zero. The velocity and pressure v_1, p_1, for the fluid are

$$\left.\begin{array}{l} v_1 = v_0 \\ p_1 = p_0 \end{array}\right\} \text{ for } z < ct \qquad \left.\begin{array}{l} v_1 = 0 \\ p_1 = 0 \end{array}\right\} \text{ for } z > ct \quad \ldots\ldots [9]$$

The second part is the response of the tube and fluid to a progressive load pressure p_2 equal to the pressure produced by the shock wave of the first part. This load pressure acts on the tube wall radially outward. Its magnitude is p_0 for $0 < z < ct$ and zero elsewhere.

Inasmuch as w_1, u_1, v_1, etc., are known, the problem reduces to finding the response w_2, u_2, v_2, etc., of the tube and fluid to the load pressure p_2. For this second part of the solution the system is initially at rest and the governing equations are Equations [4], [5], [6], [7] with

$$q_2 = \rho_0\left[\frac{\partial\phi}{\partial t}\right]_{r=a} + p_2 \quad \ldots\ldots\ldots\ldots [10]$$

The solution is begun by transforming the governing equations and eliminating the velocity potential ϕ_2. The remaining transformed equations are solved simultaneously and the inversions of the transformations then yield the solution.

With respect to z a Fourier transform[3] is used. Defining the transform of a function $g(z, t)$ with respect to z by

$$\bar{g}(\xi, t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} g(z, t)e^{-i\xi z}\,dz \quad \ldots\ldots\ldots [11]$$

then the inversion is

$$g(z, t) = \int_{-\infty}^{\infty} \bar{g}(\xi, t)e^{i\xi z}\,d\xi \quad \ldots\ldots\ldots\ldots [12]$$

With respect to time a Laplace transform is used, but the transform variable adopted is $-i = -\sqrt{-1}$ times the usual transform variable. This results in a certain symmetry with the first transform and convenience in later analysis. Defining the transform of a function $\bar{g}(\xi, t)$ with respect to t by

$$\bar{\bar{g}}(\xi, \Omega) = \frac{1}{2\pi}\int_0^{\infty} \bar{g}(\xi, t)e^{-i\Omega t}\,dt \quad \ldots\ldots\ldots [13]$$

then the inversion is

$$\bar{g}(\xi, t) = \int_{-\infty - i\alpha}^{\infty - i\alpha} \bar{\bar{g}}(\xi, \Omega)e^{i\Omega t}d\Omega \quad \ldots\ldots\ldots [14]$$

The double transform of derivatives of $g(z, t)$ may be expressed in terms of the double transform of $g(z, t)$ by the usual partial integration procedure.[4] Using the fact that initial veloci-

[3] Reference (15), p. 42.
[4] Reference (13), p. 27.

ties and displacements are zero, the transforms of the derivatives involved are

$$\overline{\frac{\partial^n}{\partial z^n}\frac{\partial^r}{\partial t^r}[g(z,t)]} = (i\xi)^n(i\Omega)^r\bar{g}(\xi,\Omega) \quad [15]$$

Using q_2 as given by Equation [10], the double transformation of Equations [4], [5], [6], and [7] yields

$$\bar{\phi}_{2rr} + \frac{1}{r}\bar{\phi}_{2r} - \xi^2\bar{\phi}_2 = -\frac{\Omega^2}{c^2}\bar{\phi}_2 \quad [16]$$

$$-\frac{\Omega^2}{c_0^2}\bar{w}_2 - \xi^2\Omega^2\frac{h^2}{12c_0^2}\bar{w}_2 + \left(1+\frac{h^2}{12a^2}\right)\frac{\bar{w}_2}{a^2} + i\xi\frac{\nu}{a}\bar{u}_2 + \frac{\xi^2h^4}{12}\bar{w}_2 + i\xi^3\frac{h^2}{12a}\bar{u}_2 = i\Omega\frac{\rho_0}{mc_0^2}[\bar{\phi}_2]_{r=a} + \frac{\bar{p}_2}{mc_0^2} \quad [17]$$

$$-\frac{\Omega^2}{c_0^2}\bar{u}_2 + \xi^2\bar{u}_2 - i\xi\frac{\nu}{a}\bar{w}_2 - i\xi^3\frac{h^2}{12a}\bar{w}_2 = 0 \quad [18]$$

$$i\Omega\bar{w}_2 = -[\bar{\phi}_{2r}]_{r=a} \quad [19]$$

The double transforms are functions of ξ and Ω only except for $\bar{\phi}_2$ which is $\bar{\phi}_2(\xi, \Omega, r)$. From Equation [16] it is found that the r dependence of $\bar{\phi}_2$ involves a Bessel function of the first kind. The function p_2 is known. Solving Equations [16], [17], [18], [19] for $\bar{w}_2$, $\bar{u}_2$, $\bar{\phi}_2$ and inverting yields

$$w_2 = -\frac{p_0}{4\pi^2m}\int_{-\infty}^{\infty}e^{i\xi z}\,d\xi\int_{-\infty-ia}^{\infty-ia}\frac{e^{i\Omega t}}{\Omega^2\left(\frac{\Omega}{c}+\xi\right)\left[\frac{\rho_0 a}{2m}\Delta+\frac{1}{N}\right]}d\Omega \quad [20]$$

$$u_2 = \frac{p_0}{4\pi^2m}\int_{-\infty}^{\infty}e^{i\xi z}\,d\xi\int_{-\infty-ia}^{\infty-ia}\frac{\nu i\xi\left(1+\frac{h^2}{12\nu}\xi^2\right)e^{i\Omega t}}{a\left(\frac{\Omega^2}{c_0^2}-\xi^2\right)\Omega^2\left(\frac{\Omega}{c}+\xi\right)\left[\frac{\rho_0 a}{2m}\Delta+\frac{1}{N}\right]}d\Omega \quad [21]$$

$$\phi_2 = -\frac{p_0}{4\pi^2m}\int_{-\infty}^{\infty}e^{i\xi z}\,d\xi\int_{-\infty-ia}^{\infty-ia}\frac{iJ_0\left(r\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\right)e^{i\Omega t}\,d\Omega}{\Omega^2\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\,J_1\left(a\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\right)\left(\frac{\Omega}{c}+\xi\right)\left[\frac{\rho_0 a}{2m}\Delta+\frac{1}{N}\right]} \quad [22]$$

where

$$\frac{1}{N} = -1 + \frac{\Omega^2 - c_0^2(1-\nu^2)\xi^2}{\Omega^2a^2\left(\frac{\Omega^2}{c_0^2}-\xi^2\right)} + \frac{h^2}{12}\left[-\xi^2 + \frac{c_0^2}{a^4\Omega^2} + \frac{2c_0^2\nu\xi^4}{\Omega^2a^2\left(\frac{\Omega^2}{c_0^2}-\xi^2\right)} + \frac{c_0^2\xi^4}{\Omega^2} + \frac{c_0^2h^2\xi^2}{12\Omega^2a^2\left(\frac{\Omega^2}{c_0^2}-\xi^2\right)}\right] \quad [23]$$

$$\Delta = \frac{2J_0\left(a\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\right)}{a\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\,J_1\left(a\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\right)} \quad [24]$$

The evaluation of the radial deflection w_2 will be considered first. It is typical of the quantities of interest. The inner or Ω integral of Equation [20] may be evaluated by Cauchy's residue theorem applied in the complex Ω-plane. During this process ξ is regarded as a constant real parameter since the outer integral involves only real values of ξ.

The factor Ω^2 in the denominator of the integrand in Equation [20] produces a simple pole at $\Omega = 0$ because

$$\left(\Omega^2\frac{1}{N}\right)$$

is finite at $\Omega = 0$. There is no pole at $(\Omega/c + \xi) = 0$ because $(\Omega/c + \xi)\Delta$ is finite at this point. There are poles at the points for which the remaining bracket is zero, i.e., where

$$\left[\frac{\rho_0 a}{2m}\Delta + \frac{1}{N}\right] = 0 \quad [25]$$

For any real ξ this equation has an infinite number of roots Ω_n. These roots have an interesting physical significance. They are the circular frequencies of the modes of free vibration of the tube and fluid system having a wave length $\lambda = 2\pi/\xi$. Alternatively, the ratios Ω_n/ξ are the phase velocities c_n of progressive sinusoidal waves of wave length $2\pi/\xi$. These facts may be developed by considering the governing Equations [4 to 7] with q in Equation [5] equal to the fluid pressure only. A solution is assumed of the form

$$\left.\begin{aligned} w &= w'e^{i(\xi z+\Omega t)} \\ u &= u'e^{i(\xi z+\Omega t)} \\ \phi &= \phi' f(r)e^{i(\xi z+\Omega t)} \end{aligned}\right\} \quad [26]$$

where the prime quantities are constants and $f(r)$ is an unknown function of r. Substituting Equations [26] in Equations [4 to 7] the constants and $f(r)$ may be eliminated, leading to exactly the transcendental Equation [25] relating ξ and Ω.

For real values of ξ all the roots Ω_n of Equation [25] are also real. If any of the Ω_n were complex or imaginary the corresponding free vibration would grow or decay indefinitely in time. Since the system has no means of dissipating or radiating energy, this is not possible physically.

Fig. 2 shows the Ω_n versus ξ-values for a typical case in which the velocity of sound in the tube wall is greater than in the fluid, i.e., $c_0 > c$. It is found that for any value of ξ there is one root Ω_1, for which

$$\left(\frac{\Omega^2}{c^2}-\xi^2\right)$$

is negative. This root corresponds to a phase velocity less than c, the velocity of sound in the fluid. There is also an infinite number of roots for which

$$\left(\frac{\Omega^2}{c^2}-\xi^2\right)$$

is positive. These correspond to phase velocities greater than c. Fig. 3 shows the same data as in Fig. 2 plotted as phase velocity versus wave length.

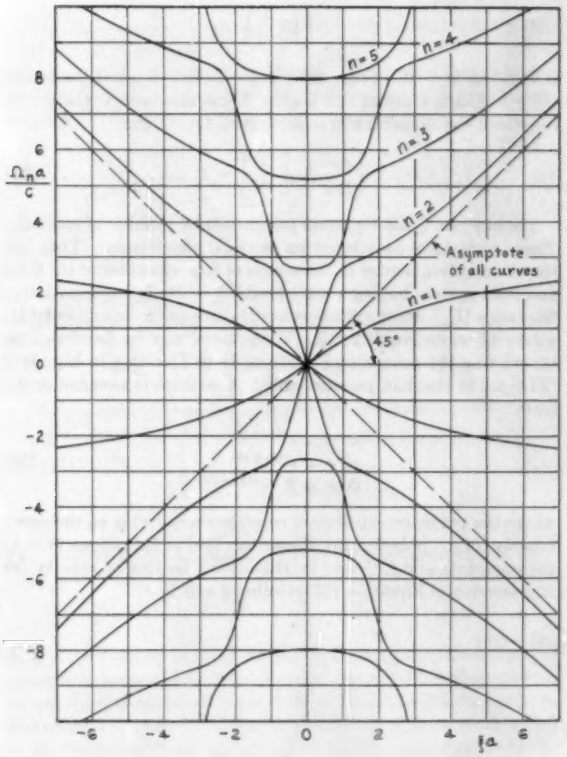


FIG. 2 FREE-VIBRATION FREQUENCIES. STEEL TUBE AND WATER $a/h = 62.8$

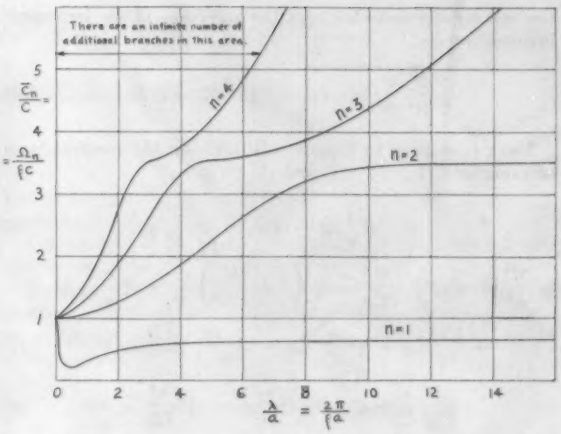


FIG. 3 PHASE VELOCITIES. STEEL TUBE AND WATER $a/h = 62.8$

The fact that two sets of Ω_n-curves pass through the origin in Fig. 2 shows that there are two modes which have finite phase velocities as the wave length approaches infinity and the frequency approaches zero. This also was shown by Lamb.[5] It is mentioned here because in a recent paper[6] it is erroneously stated that there is a cutoff frequency below which only one mode may be propagated.

The foregoing discussion shows that the poles of the integrand of Equation [20] are at $\Omega = 0$ and $\Omega = \Omega_n$, all on the real Ω-axis. The path of the integration is below this axis and it will be closed in the usual manner by adding a semicircle of infinite radius in the upper half of the Ω-plane. The integral over the semicircle approaches zero as the radius approaches infinity. The contribution of the pole at $\Omega = 0$ to w_2 is

$$w_{2a} = \frac{p_0}{2\pi i m}\int_{-\infty}^{\infty} \frac{e^{i\xi z}}{\xi\left[\frac{c_0^2}{a^2}\left(1-\nu^2+\frac{h^2}{12a^2}\right)-\frac{c_0^2\nu h^2}{6a^2}\xi^2+\frac{h^2c_0^2}{12}\left(1-\frac{h^2}{12a^2}\right)\xi^4\right]}\,d\xi \quad [27]$$

The contributions of the Ω_n-poles to w_2 may be expressed as

$$w_{2b} = -\frac{p_0}{4\pi^2 m}\int_{-\infty}^{\infty} e^{i\xi z}\,2\pi i\sum_{\Omega_n}\left[\frac{e^{i\Omega t}}{\Omega^3\left(\frac{\Omega}{c}+\xi\right)\frac{d}{d\Omega}\left[\frac{\rho_0 a}{2m}\Delta+\frac{1}{N}\right]}\right]_{\Omega=\Omega_n} d\xi \quad [28]$$

The sum indicated is over the Ω_n roots for each value of ξ. The values of the Ω_n as functions of ξ are given by Fig. 2. Equation [28] may be written in the form

$$w_{2b} = \frac{-p_0}{2\pi i m}\int_{-\infty}^{\infty}\sum_{\Omega_n}\frac{e^{i(\xi z+\Omega t)}}{\left(\frac{\Omega}{c}+\xi\right)D}\,d\xi \quad [29]$$

where

$$D = -\Omega^2\frac{d}{d\Omega}\left[\frac{\rho_0 a}{2m}\Delta+\frac{1}{N}\right] = \Omega^2\left\{\frac{\rho_0\Omega a}{mc^2\left(\frac{\Omega^2}{c^2}-\xi^2\right)}\left[1+\frac{J_0^2}{J_1^2}\right]\right.$$
$$+\frac{2C_0^2}{\Omega^2 a^2}\left[\frac{\nu^2}{\left(1-c_0^2\frac{\xi^2}{\Omega^2}\right)^2}+1-\nu^2+\frac{h^2(1+a^4\xi^4)}{12a^2}\right.$$
$$\left.\left.+\frac{h^2\xi^4\left(2\frac{\Omega^2}{c_0^2}-\xi^2\right)}{12\left(\frac{\Omega^2}{c_0^2}-\xi^2\right)^2}\left(2\nu+\frac{h^2\xi^2}{12}\right)\right]\right\} \quad [30]$$

The integral in Equation [29] covers all four quadrants of the Ω_n-curves in Fig. 2, but it may be expressed as an integral over the first quadrant values only because D is even in Ω and even in ξ and the Ω_n-curves are symmetric about both axes. Thus

$$w_{2b} = \frac{p_0}{\pi m}\int_0^{\infty}\sum_{\Omega_n}\left[\frac{\sin(\xi z-\Omega t)}{\left(\frac{\Omega}{c}-\xi\right)D}-\frac{\sin(\xi z+\Omega t)}{\left(\frac{\Omega}{c}+\xi\right)D}\right]d\xi \quad [31]$$

where the summation is over the Ω_n-curves in only the first quadrant of Fig. 2. This concludes the formal solution for w. The evaluation of the integrals remaining in Equations [27] and [31] will be considered in the next section.

[5] Reference (10), p. 7.
[6] Reference (14), p. 933.

Instead of evaluating u_2 and ϕ_2, certain quantities of interest which are derivatives of u_2 and ϕ_2 will be computed directly. These are

$$\left.\begin{aligned} &\text{Pressure in the fluid: } p_2 = \rho_0 \frac{\partial \phi_2}{\partial t} \\ &\text{Axial velocity of the fluid: } v_2 = -\frac{\partial \phi_2}{\partial t} \\ &\text{Longitudinal strain in the tube: } u_{z2} = \frac{\partial u_2}{\partial z} \end{aligned}\right\} \quad \ldots [32]$$

Expressions for p_2, v_2, and u_{z2} are derived from the double integrals for u_2 and ϕ_2, Equations [21] and [22], through application of Equation [15]. The inner integral in Ω may be evaluated in each case by using Cauchy's residue theorem as in the case of w_2. The results are

$$p_2 = \frac{p_0}{2\pi i}\int_{-\infty}^{\infty} \frac{e^{i\xi(z-ct)}}{\xi}\,d\xi + \frac{p_0\rho_0}{\pi m}\int_0^{\infty}\sum_{\Omega_n}\left[\frac{\sin(\xi z - \Omega t)}{\left(\frac{\Omega}{c}-\xi\right)F} - \frac{\sin(\xi z + \Omega t)}{\left(\frac{\Omega}{c}+\xi\right)F}\right]d\xi \quad \ldots [33]$$

$$v_2 = \frac{p_0}{2\pi i \rho_0 c}\int_{-\infty}^{\infty} \frac{e^{i\xi(z-ct)}}{\xi}\,d\xi + \frac{p_0}{\pi m}\int_0^{\infty}\sum_{\Omega_n}\left[\frac{\sin(\xi z - \Omega t)}{\left(\frac{\Omega}{c}-\xi\right)G} + \frac{\sin(\xi z + \Omega t)}{\left(\frac{\Omega}{c}+\xi\right)G}\right]d\xi \quad \ldots [34]$$

$$u_{z2} = \frac{ip_0\nu}{2\pi m a}\int_{-\infty}^{\infty}\frac{\left(1+\frac{h^2\xi^2}{12\nu}\right)e^{i\xi z}}{f(\xi)}\,d\xi + \frac{p_0\nu}{\pi m a}\int_0^{\infty}\sum_{\Omega_n}\left[\frac{\sin(\xi z - \Omega t)}{\left(\frac{\Omega}{c}-\xi\right)B} - \frac{\sin(\xi z + \Omega t)}{\left(\frac{\Omega}{c}+\xi\right)B}\right]d\xi \quad \ldots [35]$$

where

$$F = -\frac{D}{\Omega^2}\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\,\frac{J_1\left(a\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\right)}{J_0\left(r\sqrt{\frac{\Omega^2}{c^2}-\xi^2}\right)} \quad \ldots [36]$$

$$G = F\frac{\Omega}{\xi} \quad \ldots [37]$$

$$B = +\frac{D\left(\frac{\Omega^2}{c_0^2}-\xi^2\right)}{\xi^2\left(1+\frac{h^2}{12\nu}\xi^2\right)} \quad \ldots [38]$$

and $f(\xi)$ stands for the denominator of the integrand in Equation [27].

The solution of the problem originally posed at the beginning of this section is given formally by

$w = w_{2a} + w_{2b}$	Equations [27] and [31]	
$p = p_1 + p_2$	Equations [9] and [33]	$\ldots [39]$
$v = v_1 + v_2$	Equations [9] and [34]	
$u_z = u_{z2}$	Equation [35]	

Evaluation of Solution for Large Values of $|z|$ and t

In the formal solution indicated by Equations [39] the integrations with respect to ξ would be difficult to carry out in general because the Ω_n-values are known as functions of ξ only through the transcendental Equation [25]. However, some information may be obtained by considering approximations of the integrals for large values of $|z|$ and t.

Consider the part of w given by Equation [31] and, in particular, values of z and t such that

$$z = c't \quad \ldots [40]$$

where c' is an arbitrarily selected velocity. Substituting in Equation [31]

$$w_{2b} = \frac{p_0}{\pi m}\int_0^{\infty}\sum_{\Omega_n}\left[\frac{\sin(\xi c' - \Omega)t}{\left(\frac{\Omega}{c}-\xi\right)D} - \frac{\sin(\xi c' + \Omega)t}{\left(\frac{\Omega}{c}+\xi\right)D}\right]d\xi \quad \ldots [41]$$

It is found by inspection of D, that the functions

$$\frac{1}{\left(\frac{\Omega}{c}-\xi\right)D} \quad \text{and} \quad \frac{1}{\left(\frac{\Omega}{c}+\xi\right)D}$$

are bounded for all of the Ω_n-curves for all ξ except for the lowest two curves, Ω_1 and Ω_2, at the point $\xi = 0$. The distinction between the lowest two and the remaining Ω_n-curves arises because only the Ω_1 and Ω_2-curves pass through the origin. A physical interpretation of this fact is that only these two lowest modes have finite phase velocities as the wave length increases indefinitely.

For the higher Ω_n branches, $n > 2$, the Integral [41] may be approximated by the method of stationary phase,[7] for large values of t. This method shows these integrals to be of the order of $1/\sqrt{t}$.

The integrals over the Ω_1 and Ω_2-curves will be broken into two parts; one part from $\xi = 0$ to $\xi = \epsilon$ and a second part from ϵ to ∞, where ϵ is a small positive number. By the stationary-phase theorem the portions from ϵ to ∞ are again $0(1/\sqrt{t})$. Collecting these results, Equation [41] becomes

$$w_{2b} = \frac{p_0}{\pi m}\int_0^{\epsilon}\sum_{\Omega_{1,2}}\left[\frac{\sin(\xi c' - \Omega)t}{\left(\frac{\Omega}{c}-\xi\right)D} - \frac{\sin(\xi c' + \Omega)t}{\left(\frac{\Omega}{c}+\xi\right)D}\right]d\xi + 0\left(\frac{1}{\sqrt{t}}\right) \quad \ldots [42]$$

Since ϵ may be chosen indefinitely small, the functions in the integrand may be approximated by an appropriate number of terms of their expansions about $\xi = 0$. Consider the first term of the integral with $\Omega = \Omega_1$ which may be written

$$w_I = \frac{p_0}{\pi m}\int_0^{\epsilon}\frac{\sin(\xi c' - \Omega_1)t}{\xi\left(\frac{\Omega_1}{\xi c}-1\right)D}\,d\xi \quad \ldots [43]$$

The denominator of the integrand approaches a constant times ξ as ξ approaches zero. The constant is

$$\lim_{\xi\to 0}\left(\frac{\Omega_1}{\xi c}-1\right)D = \left(\frac{c_1}{c}-1\right)D_1 \quad \ldots [44]$$

where

[7] Reference (7), p. 505.

$$\left.\begin{aligned} c_1 &= \lim_{\xi \to 0}\left(\frac{\Omega_1}{\xi}\right) \\ D_1 = \lim_{\xi \to 0} D(\Omega_1, \xi) &= \frac{4\rho_0 c_1^4}{mc^2 a\left(\frac{c_1^2}{c^2} - 1\right)^2} \\ &+ \frac{2c_0^2}{a^2}\left[\frac{\nu^2}{\left(1 - \frac{c_0^2}{c_1^2}\right)^2} + 1 - \nu^2 + \frac{h^2}{12a^2}\right] \end{aligned}\right\} \quad [45]$$

The constant c_1 is the phase velocity of the slowest waves for wave lengths approaching infinity. It is very close to the Joukowsky velocity c_0. An expression for c_1 is derived in the Appendix.

In the numerator of Equation [43] Ω_1 is replaced by the first two terms of its expansion about $\xi = 0$

$$\Omega_1 = c_1\xi - d_1\xi^3 \quad [46]$$

This expansion is developed in the Appendix. It is necessary to use two terms of the expansion because the first term alone would result in a vanishing numerator for the case $c' = c_1$.

Using Equations [44] and [46] in [43] yields

$$w_I = \frac{p_0}{\pi m\left(\frac{c_1}{c} - 1\right) D_1}\int_0^\infty \frac{\sin(\xi z' + d_1\xi^3 t)}{\xi}\, d\xi \quad [47]$$

where z' is defined by

$$z' = (c' - c_1)t \quad [48]$$

The new co-ordinate z' is the location of any point relative to an origin moving at velocity c_1. In Equation [47] the upper limit has been changed from ϵ to ∞. The integral thus added is $0(1/\sqrt{t})$ by the stationary phase theorem, and is negligible in the present approximation.

For $z' = 0$, the integral in Equation [47] has a constant value of $\pi/6$ provided d_1 is positive which is the case here. For points where $|z'|$ is large, the term containing ξ^3 may be neglected. Then the integral in Equation [47] becomes $\pm\pi/2$ for $\pm z'$. For points near $z' = 0$ it is convenient to change the variable of integration from ξ to $\eta = \xi z'$. Then

$$w_I = \frac{p_0}{\pi m\left(\frac{c_1}{c} - 1\right) D_1}\int_0^\infty \frac{\sin(\eta + \beta\eta^3)}{\eta}\, d\eta \quad [49]$$

where

$$\beta = \frac{d_1 t}{z'^3} \quad [50]$$

From Equation [49] it may be seen that for w_I constant β must be constant, say, β_0. Then

$$z' = \sqrt[3]{d_1 t \beta_0} \quad [51]$$

This shows that points for which w_I is constant continually move away from the point $z' = 0$. To summarize, w_I is a wave having constant deflection at the point $z' = 0$, i.e., moving with velocity c_1. At points far ahead and far behind this wave-front point the deflection is also constant, but the wave front continually spreads out around the point $z' = 0$. Fig. 4 shows the shape of a typical wave front. In addition to the sloping front there is an oscillation for $z' < 0$ which asymptotically approaches the final value. These oscillations are due to the inclusion of the radial inertia of the system which causes an overshoot above the final value. When seeking experimental confirmation, it must be realized that these oscillations and the overshoot will be significant only when the original disturbance has a very steep front.

If the same reasoning that has been used to evaluate w_I is applied to the remaining terms of Equation [42] similar results are obtained. Thus for large t

$$w_{3b} = I_1 - H_1 + I_2 - H_2 \quad [52]$$

where

$$\frac{I_n}{H_n} = \frac{p_0}{\pi m\left(\frac{c_n}{c} \mp 1\right) D_n}\int_0^\infty \frac{\sin(z \mp c_n t \pm d_n\xi^3 t)\,\xi}{\xi}\, d\xi \quad [53]$$

The constants c_2, D_2, d_2 which appear in H_2 and I_2 are defined in the same way as c_1, D_1, d_1 in Equations [45] and [46] by changing all subscripts from 1 to 2 in these equations. The velocity c_2 is very nearly the velocity of sound in the tube wall, c_0.

The term I_1 is the deflection w_I discussed in the foregoing. The H_1-term is a wave of the same velocity and shape as w_I but of smaller amplitude and traveling in the negative direction.

The H_2 and I_2 terms are also waves of the same shape as w_I, but traveling at the velocity c_2. These precursor waves are found to be very small compared to the main water-hammer waves I_1, H_1 which travel at velocity c_1.

The complete deflection w is the sum of w_{3b}, Equation [52], and w_{3a}, Equation [27]. The deflection w_{3a} is independent of time

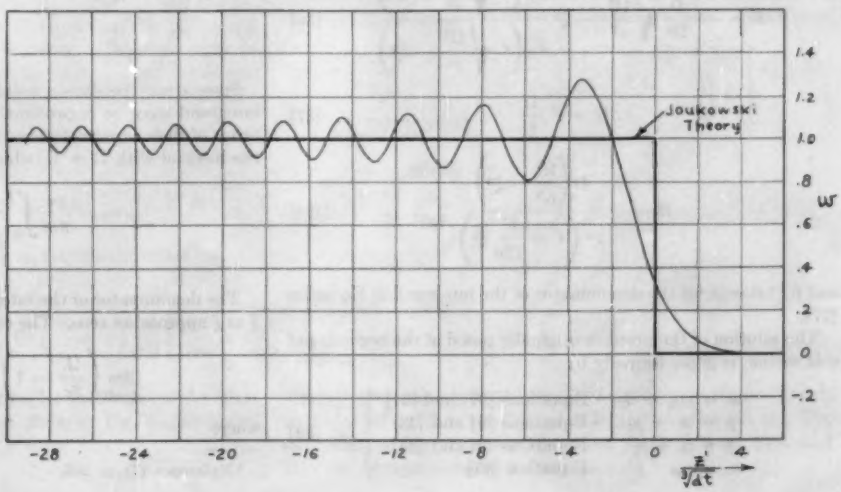


FIG. 4 FORM OF TYPICAL WAVE FRONT

and its value for large values of $|s|$ may be approximated by the same kind of reasoning as used for w_{2b} for large values of t. The result is

$$w_{2a} = \pm \frac{p_0 a^2}{2Eh\left[1 + \frac{h^2}{12a^2(1-\nu^2)}\right]} \text{ for } \pm z \quad \ldots\ldots [54]$$

The term

$$\frac{h^2}{12a^2(1-\nu^2)}$$

will be neglected compared to unity since only thin-walled tubes are considered. Thus

$$w_{2a} = \pm \frac{p_0 a^2}{2Eh} \text{ for } \pm z \quad \ldots\ldots [55]$$

The procedures used to evaluate w also may be used to obtain approximations of p, v, and u_s for large $|z|$ and t starting with the general expressions, Equations [39]. The results are

$$\left.\begin{aligned}
w &= \pm \frac{p_0 a^2}{2Eh} + I_1 - H_1 + I_2 - H_2 \quad \text{for } \pm z \\
p &= \frac{p_0}{2} - \frac{2\rho_0 c_1^2}{a\left[\frac{c_1^2}{c^2} - 1\right]}(I_1 - H_1) \\
&\quad - \frac{2\rho_0 c_2^2}{a\left[\frac{c_2^2}{c^2} - 1\right]}(I_2 - H_2) \\
v &= \frac{p_0}{2\rho_0 c} - \frac{2c_1}{a\left[\frac{c_1^2}{c^2} - 1\right]}(I_1 + H_1) \\
&\quad - \frac{2c_2}{a\left[\frac{c_2^2}{c^2} - 1\right]}(I_2 + H_2) \\
u_s &= \mp \frac{p_0 \nu a}{2Eh} + \frac{\nu}{a\left[\frac{c_1^2}{c_0^2} - 1\right]}(I_1 - H_1) \\
&\quad + \frac{\nu}{a\left[\frac{c_2^2}{c_0^2} - 1\right]}(I_2 - H_2) \text{ for } \pm z
\end{aligned}\right\} \ldots [56]$$

These equations lead to four wave fronts in each case as discussed for w. Numerical results for a typical case of water in a thin steel tube are shown in Fig. 5. In the present approximation for large t, the pressure and fluid velocity are again independent of r; i.e., they are constant across any section.

In order to estimate the length over which a given wave front extends, the slope w_s will be computed at the points which move with constant velocity c_1 or c_2. The difference in w before and after a wave front divided by the slope at the wave-front point will be indicative of the length of a wave front.

The double integral for w_s is derived from w, Equation [20]. The first integral may be evaluated and the result folded by the same procedures used for w. For a point moving so that $z = c_1 t$

$$w_s = \frac{p_0}{\pi m}\int_0^\infty \frac{\xi \cos \xi c_1 t}{f(\xi)}\,d\xi + \frac{p_0}{\pi m}\int_0^\infty \sum_{\Omega_s}\left[\frac{\xi \cos (c_1\xi - \Omega)t}{\left(\frac{\Omega}{c} - \xi\right)D} - \frac{\xi \cos (c_1\xi + \Omega)t}{\left(\frac{\Omega}{c} + \xi\right)D}\right]d\xi \quad \ldots\ldots [57]$$

The function

$$\frac{\xi}{f(\xi)}$$

is bounded for all ξ and the first integral in Equation [57] is $0(1/t)$ by the Riemann-Lebesgue lemmas.[8]

In the second integral the integrand also is bounded. This integral may be approximated by the method of stationary phase. The dominant part of the integral will come from the vicinity of stationary-phase points which are defined by

$$\frac{d}{d\xi}(\xi c_1 \pm \Omega) = 0, \text{ i.e., } \frac{d\Omega}{d\xi} = \mp c_1 \quad \ldots\ldots [58]$$

In general, such points give a contribution which is $0(1/\sqrt{t})$, but if

$$\frac{d^2}{d\xi^2}(\xi c_1 \pm \Omega) = 0, \text{ i.e., } \frac{d^2\Omega}{d\xi^2} = 0 \quad \ldots\ldots [59]$$

then the contribution to the integral from the vicinity of this point is $0\ (1/\sqrt[3]{t})$.[9] Equations [58] and [59] are both satisfied on the Ω_1-curve at $\xi = 0$. Hence this contribution will predominate over all others for large values of t.

To evaluate the dominant term, the expression for Ω_1, Equation [46], is substituted in Equation [57]. Hence for large t

$$w_s = \frac{p_0}{\pi m}\int_0^\infty \frac{\cos (d_1\xi^3 t)}{\left(\frac{c_1}{c} - 1\right)D_1}\,d\xi \quad \ldots\ldots [60]$$

Changing the variable of integration to $\eta = \xi(\sqrt[3]{d_1 t})$

$$w_s = \frac{p_0}{\pi m\left(\frac{c_1}{c} - 1\right)D_1\sqrt[3]{d_1 t}}\int_0^\infty \cos \eta^3 d\eta \quad \ldots\ldots [61]$$

This Integral [61] is known.[10] Hence

$$w_s = \frac{p_0 \Gamma\left(\frac{1}{3}\right)\sin\frac{\pi}{3}}{3\pi m\left(\frac{c_1}{c} - 1\right)D_1\sqrt[3]{d_1 t}} \quad \ldots\ldots [62]$$

Now the change of w from a location far behind the point $z = c_1 t$ to a location far ahead of the point $z = c_1 t$ is by Equation [49]

$$\Delta w = \frac{p_0}{m\left(\frac{c_1}{c} - 1\right)D_1} \quad \ldots\ldots [63]$$

Defining the "nominal length" L_1 of the wave front to be Δw divided by w_s

$$L_1 = \frac{3\pi\sqrt[3]{d_1 t}}{\Gamma\left(\frac{1}{3}\right)\sin\frac{\pi}{3}} \quad \ldots\ldots [64]$$

Computation of the nominal length L_2 of the wave front moving at velocity c_2 results in the same formula with d_1 replaced by d_2.

Some numerical values of L_1 and L_2 are given in Table 1. It may be noted that although the length of the wave front is long

[8] Reference (10), p. 172.
[9] Reference (5), p. 40.
[10] Ibid., p. 41.

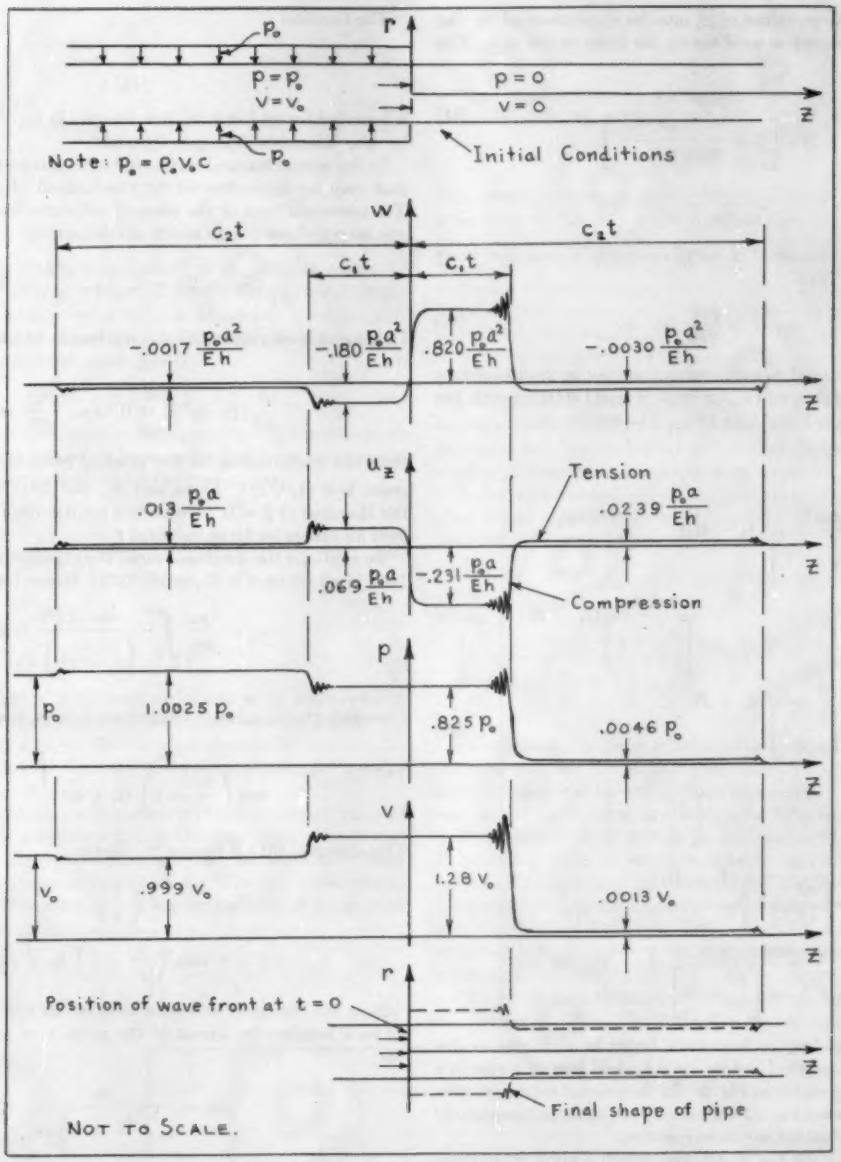


FIG. 5 VALUES OF w, u_z, p, AND v FOR A STEEL TUBE FILLED WITH WATER, $a/h = 62.8$. CASE INITIATION OF FLOW

TABLE 1 NOMINAL LENGTHS OF WAVE FRONTS FOR TYPICAL EXAMPLE

	Time, sec	Nominal length of wave front, ft	Distance of wave front from origin, ft
Primary wave	1	19.1	3218
	5	32.6	16091
Precursor wave	1	30.0	17328
	5	51.3	86640

Assumed values	Derived values
Radius of pipe: $a = 1$ ft	$a/h = 62.752$
Velocity of sound in water: $c = 5000$ fps	$R = 12.478$
Density of water: $\rho_0 = 1.94$ slugs/ft³	$c_1/c = 0.64365$
Modulus of elasticity of steel: $E = 30 \times 10^6$ psi	$c_2/c = 3.4656$
Poisson's ratio for steel: $\nu = 0.30$	$c_3/c = 3.5324$
Specific weight of steel: $\gamma_s = 490$ pcf	$c_0/c = 0.64429$
Ratio, $A = \frac{\rho_0 a}{\rho_s h}$: $A = 8$	

compared to the diameter of the pipe, it is short compared to the distance from the origin that the wave has traveled.

The foregoing computations of the lengths of wave fronts hold also for p, v, and u_z because they are described by the same integrals I_1, H_1, I_2, H_2.

AN EXTENSION OF JOUKOWSKY'S METHOD

Some of the results derived by use of integral transforms also may be derived by a much simpler procedure which is an extension of Joukowsky's method. The extension consists of taking into account the longitudinal stresses and longitudinal inertia of the pipe wall. This makes possible the precursor-type wave which is essentially a tension wave in the pipe wall.

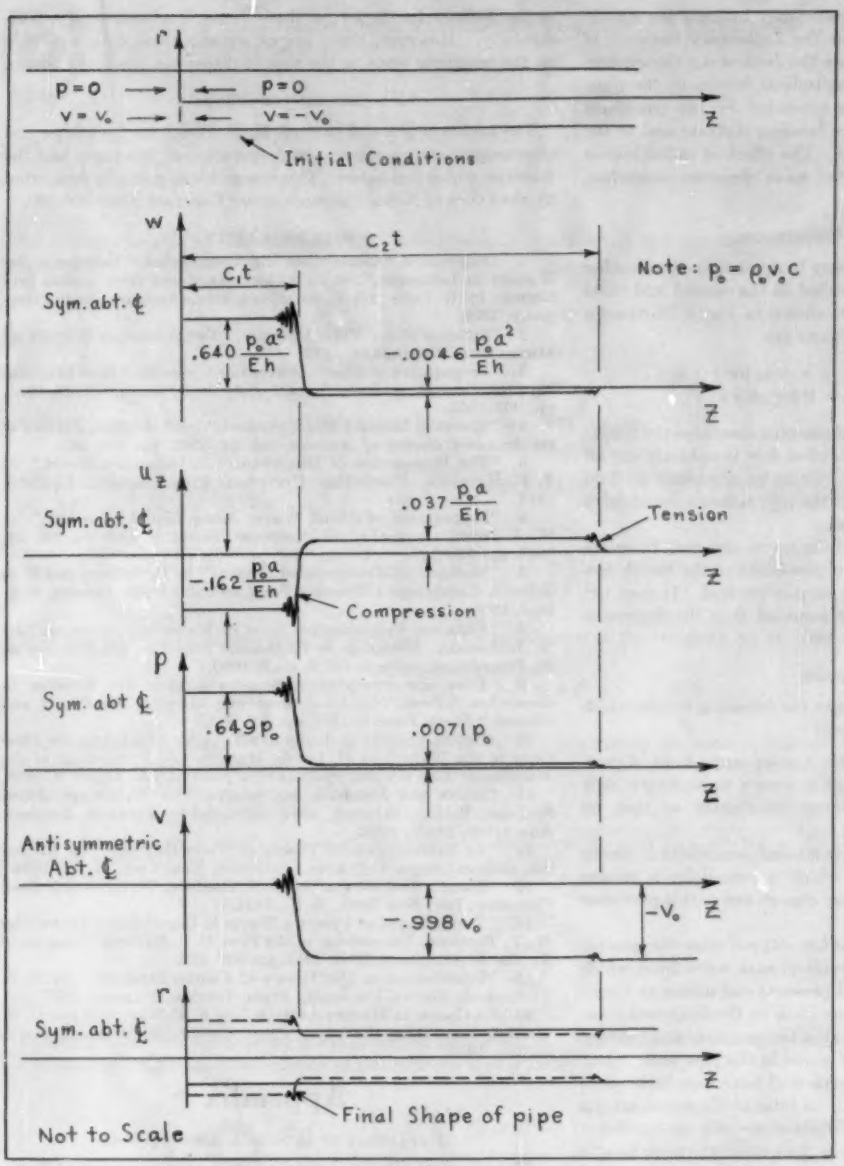


FIG. 6 VALUES OF w, u_z, p, AND v FOR A STEEL TUBE FILLED WITH WATER, $a/h = 62.8$. CASE OF SUDDEN STOPPAGE OF FLOW

The following derivation follows the general plan of Joukowsky's paper (8). The pressure and axial velocity of the fluid are assumed to be uniform over any cross section of the pipe. The fluid is assumed compressible, its bulk modulus being K. The equation of continuity may then be written in the form

$$-\frac{\partial v}{\partial z} = \frac{1}{K}\frac{\partial p}{\partial t} + \frac{2}{a}\frac{\partial w}{\partial z} \quad \ldots\ldots [65]$$

The equations of motion of the fluid reduce to

$$\rho_0 \frac{\partial v}{\partial t} = -\frac{\partial p}{\partial z} \quad \ldots\ldots [66]$$

So far, these are the equations usually used. The new element is the use of the equations of motion of the tube considered as an elastic membrane instead of a ring of no mass. These equations are derived from Equations [5] and [7] by dropping all terms arising from bending stiffness and rotatory inertia and substituting p for q

$$\frac{w}{a^2} + \frac{\nu}{a}u_z = \frac{p}{mc_0^2} \quad \ldots [67]$$

$$\frac{\ddot{u}}{c_0^2} - u_{zz} - \frac{\nu}{a}w_z = 0 \quad \ldots [68]$$

In Equation [67] the term $\ddot{w}$ has been omitted because the effect of radial inertia of the tube and of the fluid is neglected as in the Joukowsky theory.

No attempt will be made to solve Equations [65], [66], [67], and [68] in general. It is sufficient for the purpose at hand to show that the system permits solutions which are waves of arbitrary shape moving at either of the velocities c_1 or c_2 without dispersion. A solution is assumed of the form

$$\left.\begin{aligned} w &= w'f(z + \bar{c}t) \\ p &= p'f(z + \bar{c}t) \\ v &= v'f(z + \bar{c}t) \\ u_z &= u_z'f(z + \bar{c}t) \\ u_t &= u_t'f(z + \bar{c}t) \end{aligned}\right\} \quad \ldots [69]$$

where the primed quantities are constants and $\bar{c}$ is an unknown constant velocity. The constants u_t', u_z', and $\bar{c}$ are related by the requirement

$$\frac{d}{dt}u_z = \frac{d}{dz}u_t \quad \ldots [70]$$

Substituting the assumed form, Equation [69], into Equations [65], [66], [67], [68], and [70], there result five linear homogeneous equations involving the five primed constants. In order that these equations have a solution different from zero it is necessary for the determinant of the coefficients to be zero. This condition yields an equation involving $\bar{c}$ which can be satisfied only if $\bar{c}$ takes on the values $\pm c_1$ or $\pm c_2$. Thus the velocities at which waves may travel without dispersion in the present consideration are the same as the principal velocities derived in the integral-transform method.

A further point of interest is that the ratio of the change in w to the change in p across a wave front is the same in the extended Joukowsky theory and the transform solution. The same is true of any two of the quantities w, p, v, u_z, etc. This is shown by solving for the ratios of interest in each theory with $\bar{c} = c_1$ or c_2 as the case may be.

These results are of interest because they indicate the significance of the assumptions made in the Joukowsky theory. It shows that c_1 differs from c_o because the Joukowsky theory neglects longitudinal stresses and longitudinal inertia of the pipe. It also shows that the dispersion predicted by the transform method is due to the effects of the bending stiffness and of the radial inertia of the pipe and fluid. The effect of radial inertia is much more important. After the wave disperses somewhat, bending effects are negligible.

Other Initial Conditions

The integral-transform method may be applied to certain other initial conditions beside those assumed in the second and third sections. The problem and results shown in Fig. 6 illustrate a second example. The initial conditions are

$$v = v_0 \text{ for } z < 0 \qquad v = -v_0 \text{ for } z > 0$$
$$p = u = w = u_i = w_i = 0 \text{ for all } z$$

Under these conditions each half of the pipe simulates the water-hammer problem in which an established flow is suddenly cut off by an instantaneous valve closure. Since by symmetry no fluid crosses the plane $z = 0$, each half of the pipe acts as a closed valve as far as the other half is concerned.

The magnitudes and velocities of the waves also may be established on the basis of the simplified procedure of the fourth section. In general, this will be the simpler method. It does not predict dispersion, but it may be assumed that the dispersion characteristics are the same in any case.

Conclusions

The theory developed herein shows the following results which are not contained in the usual theory:

1 If the initial conditions specify a sharp wave front of pressure, this front gradually will disperse over a finite length as it proceeds. This dispersion continues indefinitely so that no steady wave shape is ever reached.

2 Two waves will be generated in general; one which is similar to the Joukowsky result and one which is essentially a tension wave in the pipe wall. The pressure change due to this precursor is very small.

3 After a sufficiently long time has elapsed since the generation of a wave there will be one point in each wave front which has a constant deflection and fluid pressure and moves at a constant velocity. This velocity is very close to the Joukowsky velocity for the main pressure wave. For the precursor this velocity is slightly less than the velocity of sound in the pipe wall.

4 The bending stresses in the pipe wall have very little effect on the characteristic velocities and the rates of dispersion after a sufficient time has elapsed. Such stresses can only be significant in the early stages of the motion in regions where the wave front is sharp so that the variation of pressure and deflection is very rapid along the length of the pipe.

5 The relations between fluid velocity, pressure, and hoop stress given by the Joukowsky theory are substantially correct, but there are also longitudinal stresses in the pipe not predicted by the usual theory.

6 Available experimental data (2) are in very close agreement with the predictions of the Joukowsky theory as to the pressure rise and velocity of the main pressure wave. Inasmuch as the analysis presented in this paper gives essentially the same results as the Joukowsky theory for these items, it is confirmed experimentally. However, there are no experimental data available on the precursor wave or the rate of dispersion predicted herein.

Acknowledgments

The author is grateful to Prof. H. H. Bleich for his advice and encouragement throughout the preparation of this paper and the thesis on which it is based. This research was partially supported by the Office of Naval Research under Contract Nonr 266(08).

Bibliography

1 "Allgemeine theorie über die veranderliche Bewegung des Wassers in Leitungen," by L. Allievi, translated from Italian into German by R. Dubs and V. Bataillard, Julius Springer, Berlin, Germany, 1909.

2 "Symposium on Water Hammer," The American Society of Mechanical Engineers, 1933.

3 "Propagation of Elastic Waves in a Cylindrical Bore Containing Fluid," by M. A. Biot, *Journal of Applied Physics*, vol. 23, 1952, pp. 997–1055.

4 "Waves in Liquid-Filled Cylinders," by R. D. Fay, *Journal of the Acoustical Society of America*, vol. 24, 1952, pp. 459–462.

5 "The Propagation of Disturbances in Dispersive Media," by T. H. Havelock, Cambridge University Press, London, England, 1914.

6 "Propagation of Sound Waves Along Liquid Cylinders," by W. J. Jacobi, *Journal of the Acoustical Society of America*, vol. 21, 1949, pp. 120–127.

7 "Methods of Mathematical Physics," by H. Jeffreys and B. S. Jeffreys, Cambridge University Press, second edition, London, England, 1950.

8 "Über den hydraulischen Stoss in Wasserleitungsrohren," by N. Joukowsky, Memoires de l'Academie Imperiale des Sciences de St. Petersbourg, series 8, vol. 9, no. 5, 1900.

9 "Über die Fortpflanzungsgeschwindigkeit des Schalles in elastisches Rohren," by D. J. Korteweg, *Annalen der Physik und Chemie*, 9 Floge, Band 5, 1878, pp. 525–542.

10 "On the Velocity of Sound in a Tube, as Affected by the Elasticity of the Walls," by H. Lamb, Memoirs and Proceedings of the Manchester Literary and Philosophical Society, vol. 42, no. 9, 1898.

11 "Statik und Dynamik der Schalen," by W. Flügge, Julius Springer, Berlin, Germany, 1934, reprinted by Edwards Brothers, Ann Arbor, Mich., 1943.

12 "An Extension of the Theory of Water Hammer," by R. Skalak, doctoral thesis, Columbia University, New York, N. Y., 1954.

13 "Fourier Transforms," by I. N. Sneddon, McGraw-Hill Book Company, Inc., New York, N. Y., 1951.

14 "Transmission of Pressure Waves in Liquid Filled Tubes," by W. T. Thomson, Proceedings of the First U. S. National Congress of Applied Mechanics, ASME, 1951, pp. 927–933.

15 "Introduction to the Theory of Fourier Integrals," by E. C. Titchmarsh, Oxford University Press, London, England, 1937.

16 "A Course in Modern Analysis," by E. T. Whittaker and G. N. Watson, Cambridge University Press, Cambridge, England, fourth edition, 1952.

Appendix

Expansion of Ω_1 and Ω_2 About $\xi = 0$

The values Ω_1, Ω_2 are defined as the lowest two roots of Equation [25]. From the numerical results represented by Fig. 2, it is expected that both Ω_1 and Ω_2 approach zero as ξ approaches zero. Hence Ω_1 and Ω_2 have an expansion of the form

$$\Omega = c\xi - d\xi^3 + \ldots\ldots\ldots\ldots\ldots\ldots [71]$$

where only odd powers occur because the Ω_n are odd functions of ξ. The values of c and d are determined by substituting this expansion in Equation [25] and equating the coefficient of each power of ξ appearing to zero.

The terms $1/N$ and Δ in the transcendental Equation [25] are given in full by Equations [23] and [24]. Expanding the Bessel functions, Δ may be written in the form

$$\Delta = \frac{\left[1 - \frac{a^2}{4}\left(\frac{\Omega^2}{c^2} - \xi^2\right) + \frac{a^4}{64}\left(\frac{\Omega^2}{c^2} - \xi^2\right)^2 + \ldots.\right]}{\frac{a^2}{4}\left(\frac{\Omega^2}{c^2} - \xi^2\right)\left[1 - \frac{a^2}{8}\left(\frac{\Omega^2}{c^2} - \xi^2\right) + \frac{a^4}{192}\left(\frac{\Omega^2}{c_0^2} - \xi^2\right)^2 + \ldots.\right]} \quad [72]$$

In order to be able to clear Equation [25] of fractions, the quotient of the two infinite series in Equation [72] is written as a single series

$$\Delta = \frac{1}{\frac{a^2}{4}\left(\frac{\Omega^2}{c^2} - \xi^2\right)}\left[1 - \frac{a^2}{8}\left(\frac{\Omega^2}{c^2} - \xi^2\right) - \frac{a^4}{192}\left(\frac{\Omega^2}{c^2} - \xi^2\right)^2 + \ldots.\right] \quad [73]$$

Using this value the transcendental Equation [25] may be put in the form

$$\frac{2\rho_0\Omega^2}{ma}\left(\frac{\Omega^2}{c_0^2} - \xi^2\right)\left[1 - \frac{a^2}{8}\left(\frac{\Omega^2}{c^2} - \xi^2\right) - \frac{a^4}{192}\left(\frac{\Omega^2}{c^2} - \xi^2\right)^2 + \ldots.\right] + \Omega^2\left(\frac{\Omega^2}{c^2} - \xi^2\right)\left(\frac{\Omega^2}{c_0^2} - \xi^2\right)\left[-1 - \frac{h^2\xi^2}{12}\right]$$

$$+ \left(\frac{\Omega^2}{c^2} - \xi^2\right)\left(\frac{\Omega^2}{c_0^2} - \xi^2\right)\frac{h^2c_0^2}{12}\left[\frac{1}{a^4} + \xi^4\right] + \left(\frac{\Omega^2}{c^2} - \xi^2\right)\frac{1}{a^2}\left[\Omega^2 - c_0^2 - c_0^2(1 - \nu^2)\xi^2 + \frac{\nu h^2 c_0^2\xi^4}{6} + \frac{c_0^2h^4\xi^6}{144}\right] = 0 \quad [74]$$

Now the assumed expansion given by Equation [71] is substituted in the foregoing for Ω. The result is a series of the form

$$b_1\xi^4 + b_2\xi^6 + b_3\xi^8 + \ldots. = 0 \quad [75]$$

Each coefficient b_n of this series must be zero since the value Ω is a root of the transcendental Equation [25]. Equating b_1 to zero and neglecting h^2/a^2 compared to unity yields an equation involving $\bar{c}$ but not d. This is interpreted as an equation on the yet unknown $\bar{c}$. The roots are

$$c_1, c_2 = c\left[\frac{2AR + R + R^2(1 - \nu^2) \mp \sqrt{[2AR + R + R^2(1 - \nu^2)]^2 - 4R^2(1 - \nu^2)(2A + R)}}{2(2A + R)}\right] \quad [76]$$

where

$$R = \frac{c_0^2}{c^2}, \qquad A = \frac{\rho_0 a}{m} \quad [77]$$

Equating b_2 to zero yields an equation involving both $\bar{c}$ and d. This equation is considered to determine two values of d, one for $\bar{c} = c_1$ and one for $\bar{c} = c_2$. The solution of the equation $b_2 = 0$ for d yields

$$d = ca^2\left[\frac{(A + 4)\left[\left(\frac{\bar{c}}{c}\right)^6 - \left(\frac{\bar{c}}{c}\right)^4(1 + R) + \left(\frac{\bar{c}}{c}\right)R\right] - \left(\frac{\bar{c}}{c}\right)\frac{h^2}{12a^2}(1 + 2R^2\nu)}{-16\left(\frac{\bar{c}}{c}\right)^2(2A + R) + 8R(2A + 1) + 8R^2(1 - \nu^2)}\right] \quad [78]$$

Discussion

JOHN PARMAKIAN.[11] This paper presents a "closer look" at the principles underlying basic water-hammer theory. For practical purposes it confirms previously established theory which has been verified experimentally on many occasions. However, it indicates some interesting possibilities with reference to very fast precursor waves in the pipe wall which may have a very minor effect on the pressure changes inside the pipe. It also estimates the length over which a sharp water-hammer wave front may extend as it moves along the axis of the pipe.

For the typical example shown in the paper at a distance of about 3 miles from the origin of the initial sharp wave form, the entire length of the extended wave front will pass a fixed point in the pipe line in about 0.01 sec. Therefore, in order to detect changes in the wave-front form, several very high-speed pressure pickups would be required at stations located many miles apart.

During the testing of the Grand Coulee Pumping Plant several years ago the writer had an opportunity to study high-frequency pressure-wave forms at two stations in the pump-discharge line.

[11] Head, Technical Engineering Analysis Section, Design Division, U. S. Bureau of Reclamation, Denver, Colo. Mem. ASME.

These stations were located about 640 ft apart. The source of the high-frequency pressure waves was the pump which since has been modified. An oscillogram of the pressure changes recorded by the high-speed pressure pickups is shown in Fig. 7, herewith. No noticeable difference in the pressure-wave forms at these two stations could be detected. Perhaps

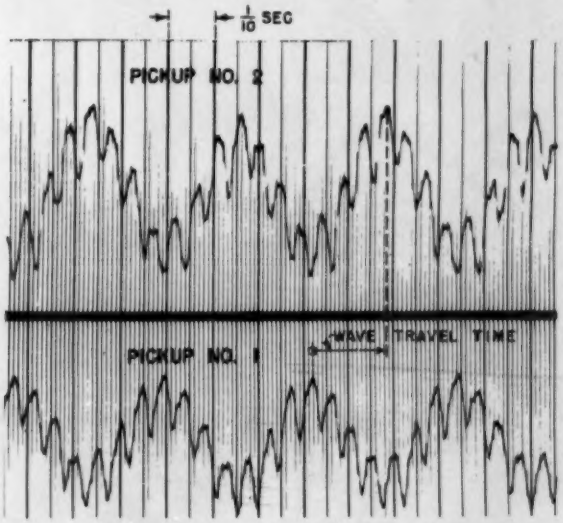


FIG. 7 OSCILLOGRAM OF PRESSURE CHANGES

the 640-ft distance between the two stations was too short and the initial wave fronts not sharp enough. In any case this experience may be indicative of the difficulty in detecting the wave-distortion phenomena of the type described in the paper.

Although the earlier water-hammer theories have been developed for a pipe consisting of individual elements which are freely extensible as noted in the paper, some of the later writers[12] have included in the theory the effect of the longitudinal restraint of the pipe line.

AUTHOR'S CLOSURE

Mr. Parmakian correctly points out that the time it takes a water-hammer wave to pass by a fixed point is very short. This is due to the fact that water-hammer waves travel at high velocities. The important physical aspect which the paper describes is the form of the wave along the pipe at a given instant. An initially sharp wave front spreads out over a considerable length in terms of pipe diameters. This change would be quite noticeable after the wave had traveled a comparatively short distance of perhaps 50 or 100 pipe diameters.

The data of Fig. 7 are interesting and are in no way contrary to the results of the paper. With reference to Fig. 3, it may be said that for any wave length greater than five times the radius of the pipe the phase velocity of the first mode is very close to the water-hammer velocity c_1. A wave length equal to 5 radii corresponds to a frequency of about 644 cps for a pipe of 1-ft radius. This means that wave forms containing only frequencies which were lower than 644 cps would be expected to disperse very slowly. The wave forms in Fig. 7 probably do not contain much energy in the frequency range above 644 cps. Therefore little dispersion should be expected. The step wave assumed in the paper contains all frequencies. Hence the dispersion due primarily to the high-frequency short-wave length components is prominent.

The reference cited by Mr. Parmakian does take longitudinal restraint into account. But it omits the effects of the inertia of the pipe wall and the radial inertia of the fluid. Therefore no dispersion of the water-hammer waves and no precursor type waves are obtained.

---

[12] See "Water-hammer Analysis," by John Parmakian, Prentice-Hall, Inc., New York, N. Y., 1955.

Design of Optimum Clearances in Positive-Displacement Pumps and Motors

By W. E. WILSON,[1] UNIVERSITY PARK, PA.

Since the efficiency of a positive-displacement pump or motor depends, to a great extent, upon the slip flow and the drag developed by moving surfaces, it is essential to proportion slip passages in such a manner that the energy losses are kept to a minimum. In this presentation, the slip flow and the drag on the moving elements are determined on the basis of a theory of laminar flow which takes into consideration the internal heating of the liquid and consequent reduction in viscosity. The optimum clearance, under various operating conditions is determined. It is shown that, under certain circumstances, the optimum clearance is not materially different from that predicted by the constant-viscosity theory. Under other circumstances, the minimum clearance that is physically possible is shown to be the optimum.

Introduction

IT IS the objective of this presentation to establish the magnitude of the clearances in positive-displacement pumps and motors which will result in the minimum loss of power due to the slip flow and the viscous drag on the moving elements. The efficiency of a pump or motor is dependent, to a major degree, on the clearances; hence one cannot emphasize too strongly the importance of the design of this feature of a positive-displacement machine.

A flow passage such as might exist at the tip of a gear tooth, the tip of a vane, or around a piston is idealized in Fig. 1. The fixed surface represents the housing, division plate, or cylinder wall as the case may be. The moving surface might be a gear or vane tip or a piston. The flow due to the pressure is defined as the slip and may be either in the direction of motion or opposed to it, depending upon the direction of the pressure gradient.

The flow between the plates due to the pressure and the motion of one plate has been described by the author.[2] This development took into consideration the heating of the liquid resulting from viscous shear and the consequent reduction in viscosity.

This theoretical description of the flow will be used here to develop the criteria for optimum clearance. It is expected that these criteria will be superior to those established previously which were based upon the assumption of constant liquid viscosity.[3] They also should be superior to those suggested earlier by the writer and based upon relatively crude efforts to include the effect of internal heating and consequent change in viscosity. The first of these assumed that the average viscosity in the flow passage was related to the clearance in quite a simple manner.[4] The second of these made use of the same theoretical background that is used in this paper and was based upon the approximation that the heat generation comes primarily from the conversion of pressure energy to heat.[5] The present paper involves neither of the previous approximations.

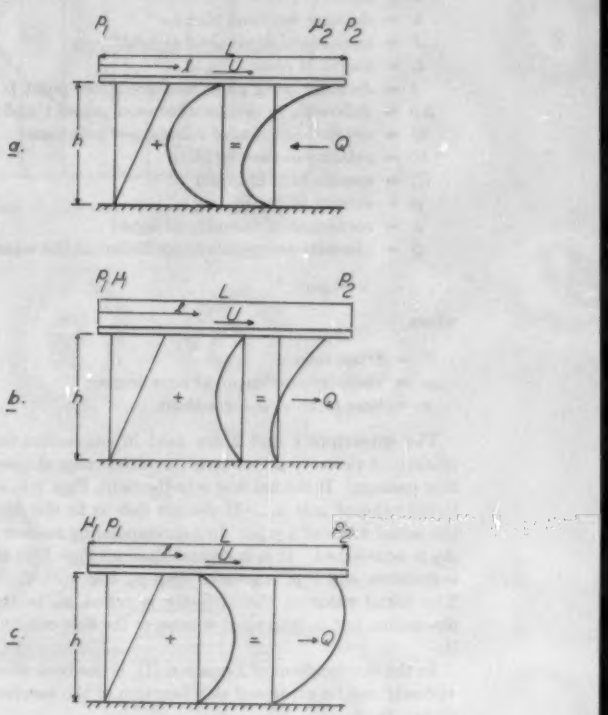


Fig. 1 Parabolic Velocity Profile

Theory

The flow between the two plates consists of the sum of two flows. The first is the shear flow, produced by the relative motion of the two plates and represented by the linear velocity profile of Fig. 1. The second is the pressure flow, produced by the difference in pressure between points 1 and 2, and represented by the parabolic velocity profile of Fig. 1. The net flow, indicated by the velocity profile, that is, the algebraic sum of the shear-flow profile and the pressure-flow profile, is in the direction of the arrow Q.

[1] George Westinghouse Professor of Engineering Education, The Pennsylvania State University. Mem. ASME.

[2] "Self-Induced Temperature Effects on Laminar Flow of Liquids," by W. E. Wilson and W. I. Mitchell, Proceedings of the First National Congress on Applied Mechanics, 1951, pp. 789–795.

[3] "Design Analysis of Rotary Pumps to Obtain Maximum Efficiency," by W. E. Wilson, *Product Engineering*, vol. 17, February, 1946, pp. 138–141. "Designing a Vane Pump," by W. E. Wilson, *Machine Design*, vol. 23, April, 1951, pp. 157–161.

Contributed by the Hydraulic Division and presented at a joint session with the Applied Mechanics Division at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–21, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, October 19, 1954. Paper No. 55—S-4.

[4] "Clearance Design in Positive Displacement Pumps," by W. E. Wilson, *Machine Design*, vol. 25, February, 1953, pp. 127–130.

[5] "Clearance Design in Hydraulic Pumps and Motors," by W. E. Wilson, *Machine Design*, vol. 26, October, 1954, pp. 249–256.

The flow is described by the following equation

$$\psi = \mp \frac{3\alpha(\alpha-1)}{1+3(\alpha-1)^2} \log_e \left\{1 \pm \frac{\theta}{\alpha}[1+3(\alpha-1)^2]\right\} \quad [1]$$

where

$$\psi = \frac{\Delta p \beta}{J\rho C_v} \quad [2]$$

$$\theta = \frac{2\beta\mu_1 UL}{J\rho C_v h^2} \quad [3]$$

$$\alpha = \frac{2Q}{Ubh} \quad [4]$$

In Equation [1], the upper sign applies to net flow to the right in Fig. 1 and the lower sign to net flow to the left. The definitions of the terms used in the dimensionless parameters θ and ψ referred to Fig. 1 are as follows:

b = width of plate
h = distance between plates
J = mechanical equivalent of heat
L = length of plate
l = distance along plate measured from point 1
Δp = difference in pressure between points 1 and 2
Q = net flow in units of volume per unit time
U = velocity of moving plate
C_v = specific heat of liquid
ρ = density of liquid
μ = coefficient of viscosity of liquid
β = viscosity-temperature coefficient in the equation

$$\mu = \mu_0 e^{-\beta T}$$

where

T = temperature
μ_0 = viscosity coefficient at zero degrees
e = base of natural logarithms

The subscripts 1 and 2 are used in connection with the coefficient of viscosity μ to denote its initial value at entrance to the flow passage. If the net flow is to the right, Figs. 1(*b*) and 1(*c*), the initial value of μ is μ_1. If the net flow is to the left, Fig. 1(*a*), the initial value of μ is μ_2. In a corresponding manner, the sign of Δp is established. If p_2 is greater than p_1, Figs. 1(*a*) and 1(*b*), Δp is positive, and if p_1 is greater than p_2, Fig. 1(*c*), Δp is negative. The initial value of the viscosity is called μ_1, in the following discussion, but μ_2 is implied whenever the flow conditions require it.

In the development of Equation [1], it has been shown that the viscosity can be expressed as a function of the location along the plate; thus

$$\mu = \frac{\mu_1}{1+\dfrac{\beta\mu_1}{QJ\rho C_v}\left[\dfrac{U^2 b}{h}+\dfrac{12}{bh^3}\left(Q-\dfrac{Ubh}{2}\right)^2\right]l}$$

Using this equation for μ and defining its average value thus

$$\mu_{avg} = \frac{\int_0^L \mu dl}{L}$$

we obtain upon integration

$$\frac{\mu_{avg}}{\mu_1} = \frac{\text{Log}_e\left\{1 \pm \dfrac{\theta}{\alpha}[1+3(\alpha-1)^2]\right\}}{\dfrac{\theta}{\alpha}[1+3(\alpha-1)^2]} \quad [5]$$

The flow described by Equations [1] and [5] is illustrated in graphical form in Figs. 2 and 3. In Fig. 2, the functional relationship among the variables α, θ, and ψ is shown. In Fig. 3, the relationship among the quantities θ, ψ, and the viscosity ratio μ_{avg}/μ_1 is shown.

Fig. 2 is best adapted to describing the general nature of the flow. The direction of the net flow is indicated in each quadrant of the chart by the arrow marked Q in the flow diagram. The flow illustrated in the upper left-hand quadrant is primarily a pressure flow to the left. The flow in the upper right-hand quadrant is net flow to the right, less in magnitude than the shear flow alone; and in the lower right-hand quadrant net flow to the right, with shear and pressure flow adding, is indicated.

The details of these situations will be discussed in connection with the viscosity ratio described in Fig. 3. In Fig. 3, the lines marked $\theta = 0$ represent a flow due to pressure only with the velocity $U = 0$, since, in a practical case, $\theta = 0$ only when U is zero. These curves indicate a decrease of viscosity with increasing pressure with an asymptotic approach to zero viscosity as the pressure increases. This flow is shown in Fig. 2 by the point $\alpha = 1$, $\psi = 0$. The curves to the left of the vertical axis of Fig. 3, which passes through the point $\mu_{avg}/\mu_1 = 0$, $\psi = 0$, describe the situation with a net flow to the right. These correspond to the curves of the lower right-hand quadrant of Fig. 2. As the pressure differential decreases, ψ decreases, the flow becomes a pure shear transport at $\psi = 0$, and its magnitude is unaffected by the viscosity, remaining constant and equal to $Ubh/2$.

The curves which lie to the right of the vertical axis describe net flow to both left and right, corresponding to flow shown in the upper left and right-hand quadrants of Fig. 2. As the positive pressure differential decreases, corresponding to movement to the left along the lower branch of the curves, a decreasing net flow is indicated. Ultimately, the net flow approaches zero owing to the balancing of shear flow to the right and pressure flow to the left. The heating of the liquid is then so great that the viscosity approaches zero. The curves are shown passing through the point

$$\psi = 0, \frac{\mu_{avg}}{\mu_1} = 0$$

In a real situation such would not be the case, since some heat would be dissipated and the viscosity would be greater than zero. The flow would then be described by the branch of the curves of Fig. 3 passing through the point 0, 0 and intersecting the vertical axis at some value of the viscosity ratio greater than zero. This corresponds to the curves of the upper right-hand quadrant of Fig. 2. This regime is one of great flow instability and represents the flow in many slip passages of pumps. The more stable flow, corresponding to negative values of ψ, occurs principally in motors. This flow is described in the lower right-hand quadrant of Fig. 2 and the portion of Fig. 3 to the left of the viscosity-ratio axis.

The total force F required to maintain the motion of the plate at the speed U can be determined by considering that

$$F = \int_0^L \mu \frac{U}{h} b dl$$

Substitution of the expression for the viscosity as a function of the location along the plate and subsequent integration result in the equation

$$F = \frac{\mu_1 UbL}{h} \frac{1}{\dfrac{\theta}{\alpha}[1+3(\alpha-1)^2]} \log_e \left\{1 \pm \frac{\theta}{\alpha}[1+3(\alpha-1)^2]\right\} \quad [6]$$

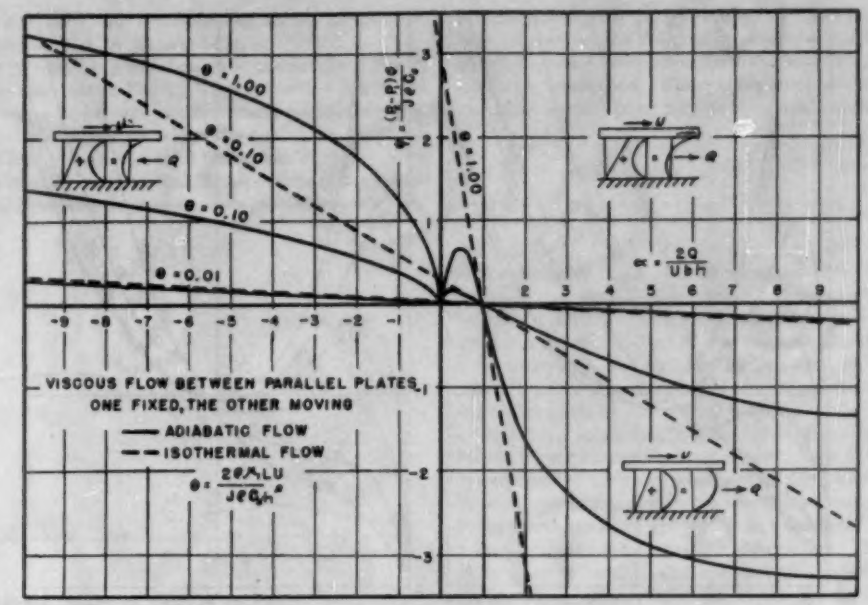


FIG. 2 VISCOUS FLOW BETWEEN PARALLEL PLATES—ONE FIXED, THE OTHER MOVING

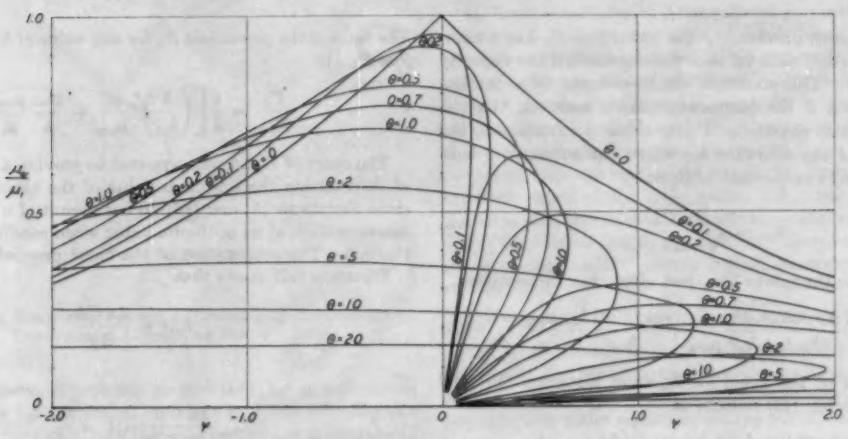


FIG. 3 VISCOSITY RATIO

which can be simplified to read

$$F = \frac{\mu_{avg} U b L}{h} \quad \ldots\ldots [7]$$

A corresponding manipulation of Equation [1] yields the following expression for the rate of flow Q between the plates

$$Q = \frac{Ubh}{2} + \frac{(p_1 - p_2) b h^3}{12 \mu_{avg} L} \quad \ldots\ldots [8]$$

The flow due to pressure represents a loss in effectiveness of the pump or motor since it results in a reduction in the delivery of the pump or an increase in the flow to the motor, compared with the ideal geometrical delivery. This flow, the slip, is the second term of Equation [8] and will be designated Q_s. We then have

$$Q_s = \frac{\Delta p b h^3}{12 \mu_{avg} L} \quad \ldots\ldots [9]$$

The drag F due to viscous friction also represents a loss of effectiveness since the energy to maintain the motion at the speed U must be supplied at the expense of the useful work of the pump or motor.

One may, therefore, write for the power loss P_L due to a single slip passage the expression

$$P_L = Q_s \Delta p + FU \quad \ldots\ldots [10]$$

where Q_s is the slip flow due to the pressure and Δp is the pressure differential $(p_2 - p_1)$. This equation may be rewritten in the following significant form

$$P_L = \frac{(\Delta p)^2 b h^3}{12 \mu_{avg} L} + \frac{\mu_{avg} U^2 b L}{h} \quad \ldots\ldots [11]$$

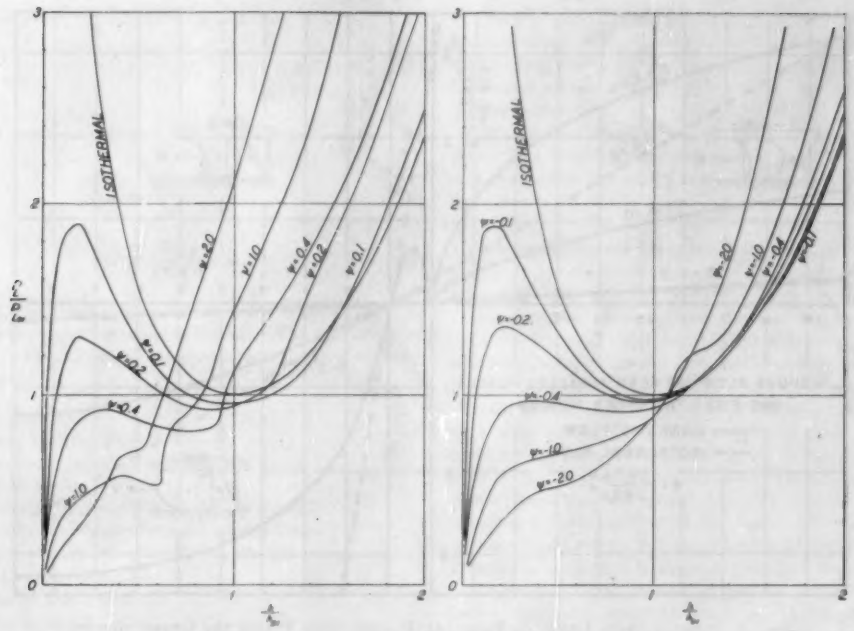


FIG. 4 POWER-LOSS RATIO

As has been shown previously,[3] the power loss P_L has a minimum value for an optimum value of the clearance, if the viscosity remains constant. This situation will be referred to as the isothermal case since, if the temperature were constant, the viscosity would remain constant. Under those circumstances, the optimum value of the clearance h_{oi} where the subscript o indicates optimum and i isothermal, is given by

$$h_{oi} = \sqrt{\frac{2\mu_1 UL}{\Delta p}} \quad \ldots\ldots\ldots\ldots\ldots\ldots [12]$$

Equation [11] for the power loss, just given, can be rewritten

$$P_L = \frac{(\Delta p)^2 b h_{oi}^3}{12\mu_1 L} \frac{h^3}{h_{oi}^3} \frac{\mu_1}{\mu_{avg}} + \frac{\mu_1 U^2 bL}{h_{oi}} \frac{h_{oi}}{h} \frac{\mu_{avg}}{\mu_1} \quad \ldots\ldots [13]$$

Now making use of the value for h_{oi} from Equation [12] and simplifying, Equation [13] reduces to

$$P_L = \frac{(\Delta p)^2 b h_{oi}^3}{12\mu_1 L} \left\{ \left(\frac{h}{h_{oi}} \right)^3 \frac{\mu_1}{\mu_{avg}} + \frac{3h_{oi}}{h} \frac{\mu_{avg}}{\mu_1} \right\} \quad \ldots\ldots [14]$$

This equation can be used to determine the optimum value of the clearance ratio h/h_{oi} by graphical means.

Fig. 4 has been prepared for this purpose. However, if the isothermal case is considered, $\mu_{avg}/\mu_1 = 1$ and the optimum value of h/h_{oi} is 1. The corresponding optimum power loss, termed P_{Loi}, is given by Equation [14] in the form

$$P_{Loi} = \frac{(\Delta p)^2 b h_{oi}^3}{12\mu_1 L} (1 + 3)$$

The value of the quantity in the bracket four is the minimum that is possible. This results in the final expression for the optimum power loss under isothermal conditions

$$P_{Loi} = \frac{(\Delta p)^2 b h_{oi}^3}{3\mu_1 L} \quad \ldots\ldots\ldots\ldots\ldots\ldots [15]$$

The ratio of the power loss P_L for any value of h/h_{oi} to the power loss P_{Loi} is

$$\frac{P_L}{P_{Loi}} = \frac{1}{4} \left[\left(\frac{h}{h_{oi}} \right)^3 \frac{\mu_1}{\mu_{avg}} + \frac{3h_{oi}}{h} \frac{\mu_{avg}}{\mu_1} \right] \quad \ldots\ldots\ldots [16]$$

The chart of Fig. 4 was prepared to provide a graphical means of determining the optimum value of the clearance ratio h/h_{oi}, since Equations [5] and [16] do not permit of a direct analytical determination of an optimum value when conditions are not isothermal. The preparation of the chart proceeded as follows.

Equation [12] shows that

$$h_{oi}^3 = \frac{2\mu_1 UL}{\Delta p}$$

or

$$\frac{2\mu_1 UL}{(\Delta p) h_{oi}^3} = 1 = \frac{\theta_{oi}}{\psi} \quad \ldots\ldots\ldots\ldots\ldots [17]$$

where θ_{oi} is the value of θ corresponding to the situation with the optimum clearance under isothermal conditions.

The definition of the parameter θ shows that

$$\theta = \frac{2\beta\mu_1 UL}{J\rho C h_{oi}^3} \left(\frac{h_{oi}}{h} \right)^3$$

or

$$\theta = \theta_{oi} \left(\frac{h_{oi}}{h} \right)^3 \quad \ldots\ldots\ldots\ldots\ldots [18]$$

The viscosity is described by Equation [5], which in turn is represented graphically in Fig. 3.

Values of the ratio h/h_{oi} and the parameter ψ were assigned. The corresponding value of θ_{oi}, equal to ψ in accordance with Equation [17], was inserted in Equation [18] and the value of θ was computed for each value of h/h_{oi}. With these values of θ and

the assigned value of ψ, the corresponding values of the ratio μ_{ave}/μ_1 were determined with the aid of Fig. 3.

It will be noted that the general shape of the curves of Fig. 3, for μ_{ave}/μ_1, is as indicated in Fig. 5. The broken line indicates a portion of the curve which does not describe a possible physical situation. Since the theory on which the curves are based assumed no heat flow through the solid boundaries or in the direction of the flow axis, it is obvious that those portions of the curves associated with extremely high temperatures cannot be wholly realistic.

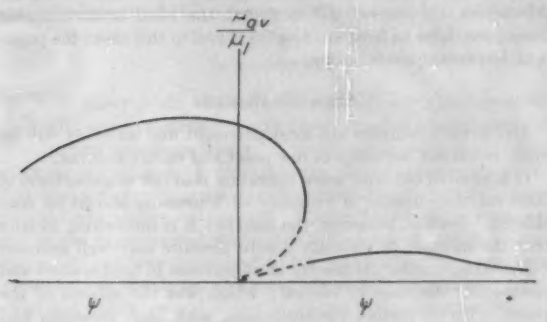


FIG. 5 GENERAL SHAPE OF CURVES OF FIG. 3 FOR μ_{ave}/μ_1

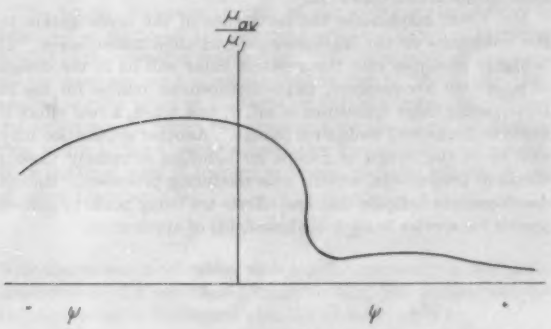


FIG. 6 GENERAL FORM FOR VERTICAL CURVES CORRESPONDING TO THEORETICAL CURVES OF FIG. 3

In Fig. 6 is shown a probable general form for actual curves corresponding to the theoretical ones in Fig. 3. In carrying out the calculations for Fig. 4, only the solid portions of the curves as indicated in Fig. 5 were used. As a consequence of this fact and the indicated instability of the flow in the region of the broken lines, one must not take too seriously the fine detail of the curves of Fig. 4 in the neighborhood of the saddle points.

The values of μ_{ave}/μ_1, shown in Fig. 3, were used in Equation [16] to determine the magnitude of the ratio P_L/P_{Loi}. In Fig. 4, this ratio is shown as a function of the clearance ratio h/h_{oi} for constant ψ. This plotting provides the basis for a rational design procedure which is described in the following section.

DESIGN PROCEDURE

The simplicity of this design method depends upon the use of the dimensionless parameters θ and ψ to express the operating conditions and the characteristics of the slip passage to be designed.

It is assumed that the physical dimensions, except the clearance, are dictated largely by considerations other than minimum power loss associated with the passage. The liquid is assumed to have been selected by the designer or otherwise specified. The speed of rotation and pressure differential also are considered to be otherwise determined. Under these circumstances, the following quantities are known or may be calculated readily:

U	J
L	μ_1
b	β
C_s	ρ
Δp	

With these, ψ and h_{oi} can be calculated directly from Equation [2] and Equation [12], respectively. Reference to Fig. 4 will then reveal the nature of the variation of the power loss with the clearance h. If the value of ψ is less than about 0.4, there exists a value of h/h_{oi}, not significantly less than unity, that yields a power loss less than that for any clearance other than extremely small clearances of the order of 0.1 h/h_{oi}.

This fact is interesting in that it confirms and gives significance to two widely divergent points of view. One of these is that held by experienced designers who have long felt that extremely small clearances would give high efficiencies and such is the case. The other was a result of the isothermal theory of clearance design which predicted the optimum clearance h_{oi}. The significance of these facts is twofold and lies first in the possibility of using comparatively large clearances of the order of h_{oi}, when very small ones would be too costly or when slightly larger losses are tolerable, and then, in the point that extremely small clearances are necessary to obtain very high efficiencies.

It should be noted that if the value of ψ exceeds 2, the optimum clearance is the least clearance physically possible. For purposes of rough estimation, a value of $\psi = 1$ corresponds to a differential pressure of approximately 5000 psi. This indicates that in the range of common practice; i.e., when Δp is less than 5000 psi, a good design can be obtained with clearances not materially less than the optimum isothermal clearance.

If very high efficiencies are sought by making clearances exceedingly small of the order of 0.1 h/h_{oi}, one may expect considerable difficulty in avoiding metal-to-metal contact and consequent large losses due to dry friction. However, extremely high pressure may require such a procedure under certain circumstances.

NONIDEAL DESIGNS

Fig. 3 is extremely useful in determining the viscous frictional drag and slip for designs that are not necessarily optimal. Such situations arise continually since pumps handle many different liquids under various operating conditions and the particular geometry is correct for one combination of circumstances only.

If all the dimensions of a slip passage are known and the operating conditions in terms of speed, pressure differential, and viscosity at the intake are known, it is then possible to compute the slip and drag.

The calculations leading to the determination of these two quantities are direct and simple. First θ and ψ must be computed as indicated previously. Using these and Fig. 3, the value of μ_{ave}/μ_1 can be found. This quantity is substituted in Equations [6] and [9] to find the drag and slip, respectively. The sum of the power losses for the individual passages will be the total loss for the pump or motor. Complete performance characteristics over a range of values of speed, pressure, and viscosity can be calculated in this manner.

SUMMARY

The design of clearances in positive-displacement pumps and motors, taking due account of the variation in the viscosity of the

liquid resulting from internal heating, can now proceed in a straightforward manner.

The limitations of the constant-viscosity theory need no longer inhibit good design.

The assumptions underlying the new theory are such that deviations from them will, in general, result in flow conditions more nearly approaching those of the isothermal case. Consequently, the two theories represent limits between which actual performance may be expected to be found.

The interpretation of test data can be facilitated particularly by means of Fig. 3. Systematic studies directed toward the evaluation of the effect of clearances on pump performance will yield data that can be analyzed in the light of facts shown in Fig. 3. Consideration of such data on the basis of the isothermal theory has been unsatisfactory since the slip flow is usually several times as great as expected and the frictional drag is correspondingly less.

Discussion

WALTER ERNST.[6] The paper deals with an important aspect of hydrostatic motor and generator design, in fact, with the basic principle of transmission of power by displacement of Newtonian fluids.

The writer, in his textbook,[7] has attempted to present the subject in somewhat more elementary form and with some simplifying assumptions, among which is the employment of the Herschel relationship between viscosity and temperature which should be adequate for the range within which the subject transmissions operate in actual practice. An extension of validity beyond that range would have little practical meaning.

The writer feels that with the great pains taken to evaluate the influence of temperature on viscosity, that of pressure on viscosity must not be neglected. At 5000-psi pressure, which is reached by some of today's commercial units, the viscosity of most hydraulic oils more than doubles.

The relationship is $\mu = \mu_0 e^{\beta P}$ with β roughly approximately 2 in the ordinary viscosity range.

There is no question concerning the didactic value of this type of investigation which greatly helps the designer in grasping the fundamentals and evaluating the influence of the variables entering into the problem.

Actual operating conditions, as the author points out, are rarely optimal, and added to this are inevitable manufacturing tolerances that may reach orders of magnitude of the clearances themselves. However, the fact that the ideal is unrealizable should not deter us from striving for it and to this effort the paper is an important contribution.

AUTHOR'S CLOSURE

Mr. Ernst's remarks are most pertinent and an effort will be made to extend the scope of one point and clarify another.

It is pointed out that when operation is in the neighborhood of 5000 psi the increase of viscosity with pressure should be considered. Such is, of course, the case but it is interesting to note that the increase of viscosity due to pressure may well counterbalance the contrary tendency of an increase in temperature and consequent decrease in viscosity which was the subject of the paper. Under certain circumstances, with high pressures and high relative velocities of adjacent surfaces, the net effect might well be that the optimum isothermal clearance would be very close to the actual optimum.

Mr. Ernst emphasizes the pertinence of the investigation to the mechanics of the high-pressure hydraulic transmission. It is highly probable that the greatest value will be in the design of relatively low-pressure, large-displacement pumps for use in transporting large quantities of oil, if, and when, a real effort is made to design and build such pumps. Another application may well be in the design of pumps for handling extremely viscous chemical products in various manufacturing processes. Recent developments indicate that real efforts are being made to perfect pumps for service in each of these fields of application.

[6] Vice-President, Director of Engineering, The Commonwealth Engineering Company of Ohio, Dayton, Ohio. Mem. ASME.

[7] "Oil Hydraulic Power and Its Industrial Applications," by Walter Ernst, McGraw-Hill Publishing Co., New York, N. Y., 1949.

Fundamentals of the Vibratory Rate Gyro

By J. B. CHATTERTON,[1] GREAT NECK, N. Y.

The important theoretical characteristics of the new vibratory type of rate gyro are discussed. Developed torque, output sensitivity, and frequency response are derived in a simple manner. While no engineering details of a practical device are mentioned, the conclusion is reached that an instrument of this type is basically sound in principle

Introduction

FOR many years the conventional rate gyro has been the only practical primary instrument for the measurement of inertial angular velocity. The vibratory rate gyro, which utilizes a vibrating rather than rotating mass, is a new instrument that also measures inertial angular velocity and appears to have several basic relative advantages. While vibratory rate gyros are still in the experimental stage, it is believed that the basic configuration[2] has been developed and is fundamental to a practical instrument of this type. It is the purpose of this paper to develop and review several of the important basic characteristics of this configuration.

Developed Torque

The torque-producing portion of the structure consists of a balanced or double-ended vibrating system, the simplest being similar to the common tuning fork. The input and output axes are the same and are located as shown in Fig. 1. With a constant applied angular velocity, Ω, Equation [29] of the Appendix shows the only force F_0, which can produce a torque about the output axis, is due to the Coriolis acceleration forces of the two masses. With the total mass M lumped as shown, the total output torque is, using Equation [29]

$$T = 2rF_0 = 2r\dot{r}M\Omega \quad \ldots\ldots [1]$$

With the masses in vibration with angular frequency ω_v and peak displacement ΔR considered small so that the motion can be assumed radial to the output axis, the distance r will be

$$r = R + \Delta R \sin \omega_v t \quad \ldots\ldots [2]$$

Substituting in Equation [1]

$$T = 2M\Delta R\omega_v\Omega\,(R + \Delta R \sin \omega_v t)\cos \omega_v t \quad \ldots\ldots [3]$$

and neglecting the small double-frequency term, the approximate torque becomes

$$T = 2MR\Delta R\omega_v\Omega \cos \omega_v t \quad \ldots\ldots [4]$$

Thus the developed torque is seen to be of an a-c nature and, since in the general case Ω will be a function of time, the torque is suppressed-carrier modulated. The carrier frequency is equal to the frequency of mass vibration.

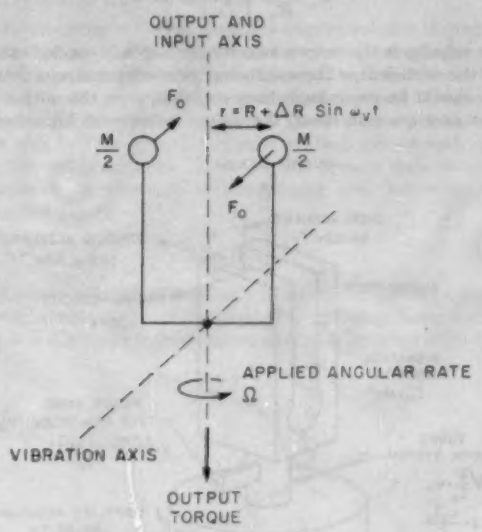


Fig. 1 Vibratory Rate-Gyro Moving Mass

Output Motion

It was tacitly assumed that the motion of the masses in the previous discussion was considered with respect to a primary or secondary inertial reference frame. It is easily seen that a particle, or system of particles such as would make up an instrument housing, also rotating with constant Ω, would have no transverse or torque-producing forces. Clearly, then, the developed output torque will act with respect to an instrument frame. As is the usual practice in conventional rate gyros, an arrangement is made where this torque will produce a torsional motion of a restraining system with respect to the instrument frame and, with an appropriate transducer, produce an output voltage or signal proportional to this motion.

The total system configuration is shown in Fig. 2. Since the output torque of Equation [4] was of a carrier type, the velocity of a spring-mass system can be utilized as the output motion and can be measured with a nondifferential, velocity-sensitive transducer. With the undamped natural frequency of the system in torsion equal to the mass-vibration frequency, or with spring constant k and total inertia about the output axis I_0

$$\omega_v = \sqrt{\frac{k}{I_0}} \quad \ldots\ldots [5]$$

and assuming this frequency is also equal to the damped natural frequency, the output velocity becomes, with damping resistance C and using Equation [4]

$$V = \frac{T}{C} = \frac{2MR\Delta R\omega_v}{C}\,\Omega \cos \omega_v t \quad \ldots\ldots [6]$$

Defining a vibration-mass inertia as

---

[1] Senior Engineer, Sperry Gyroscope Company.

[2] "A New Space Rate Sensing Instrument," by J. Lyman, and "The Gyrotron[3] Angular Rate Tachometer," by R. Barnaby, J. Chatterton, and F. Gerring, *Aeronautical Engineering Review*, vol. 12, November, 1953, pp. 24–36.

[3] Registered trade-mark U. S. Patent Office.

Contributed by the Instruments and Regulators Division and presented at a joint session of the Instruments and Regulators and Aviation Divisions at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–22, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 31, 1955. Paper No. 55—S-25.

$$I_M = MR^2 \quad [7]$$

and with the usual Q of a tuned system

$$Q = \frac{\omega_v I_0}{C} \quad [8]$$

Equation [6] becomes

$$V = \frac{2\Delta RQ}{R}\left[\frac{I_M}{I_0}\right]\Omega \cos \omega_v t \quad [9]$$

This velocity is the output velocity for a steady applied-rate input, the coefficient of $\Omega \cos \omega_v t$ being the mechanical sensitivity S.

It should be recognized, however, that since the output and input axes are coincident, the output velocity of Equation [9]

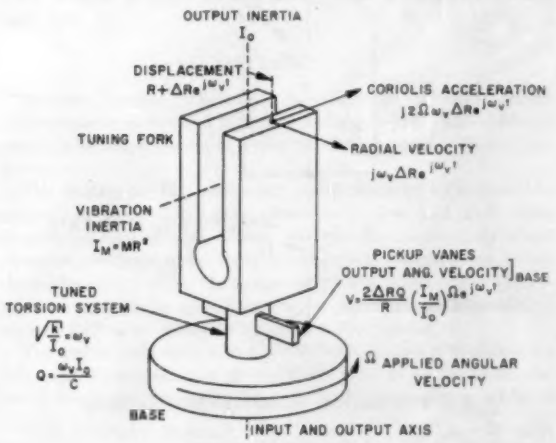


FIG. 2 VIBRATORY RATE-GYRO SYSTEM

also will be applied to each vibrating mass along with the applied velocity Ω. Because of this effect, a more accurate expression can be written for the output velocity, similar to the usual feedback equation with a feedback factor of minus unity

$$V = \frac{S \cos \omega_v t}{1 + S \cos \omega_v t}\Omega \quad [10]$$

Since the sensitivity is less than unity

$$V = S\,\Omega \cos \omega_v t\,[1 - S \cos \omega_v t + \ldots] \quad [11]$$

Retaining only the carrier frequency term results in the original velocity of Equation [9] and shows it to be a valid approximation. Also, it is seen that the output motion produces no nonlinearity of the over-all output.

OUTPUT RESPONSE

Considering the applied input velocity Ω as a function of ωt, the output frequency response could be obtained by solving for the torsion-system velocity with two applied side-band torques with respective frequencies of $(\omega_v + \omega)$ and $(\omega_v - \omega)$. More along the lines of servo analysis, however, the output response and transfer function can be deduced from the conventional response of the familiar spring-mass or second-order torsion system.

If the amplitude of the torsion system $f(\omega t)$ is assumed to be a perfect, even, or symmetrical function about the natural frequency ω_v, then for any deviation $\Delta\omega$

$$f(\omega_v t + \Delta\omega t) = f(\omega_v t - \Delta\omega t) \quad [12]$$

and if the phase $\theta(\omega)$ is odd about ω_v

$$\theta(\omega_v + \Delta\omega) = -\theta(\omega_v - \Delta\omega) \quad [13]$$

With an applied force of the form

$$T = K \cos \Delta\omega t \cos \omega_v t$$

$$= \frac{K}{2}\left[\cos(\omega_v t + \Delta\omega t) + \cos(\omega_v t - \Delta\omega t)\right] \quad [14]$$

the output of the system will be

$$V = \frac{K}{2}\{f(\omega_v t + \Delta\omega t)\cos[\omega_v t + \Delta\omega t + \theta(\omega_v + \Delta\omega)] + f(\omega_v t - \Delta\omega t)\cos[\omega_v t - \Delta\omega t + \theta(\omega_v - \Delta\omega)]\} \quad [15]$$

Rewriting and using Equations [12] and [13]

$$V = \frac{Kf(\omega_v t + \Delta\omega t)}{2}[\cos\{\omega_v t + [\Delta\omega t + \theta(\omega_v + \Delta\omega)]\} + \cos\{\omega_v t - [\Delta\omega t + \theta(\omega_v + \Delta\omega)]\}] \quad [16]$$

which is in the form of the original trigonometric transformation of Equation [14]. Therefore

$$V = Kf(\omega_v t + \Delta\omega t)\cos[\Delta\omega t + \theta(\omega_v + \Delta\omega)]\cos\omega_v t \quad [17]$$

which shows that the output response of the intelligence of the modulated wave, or rather the response of the perfectly demodulated output, is simply the response of the right-hand side of the symmetrical portion of the conventional second-order system.

The total radian bandwidth of the torsion system can be written as a function of the system Q

$$\text{BW} = \frac{\omega_v}{Q} \quad [18]$$

and one half of this value will be the frequency of applied velocity where the velocity-output amplitude will be down 3 decibels (db) from the zero frequency amplitude. The phase shift will be 45 deg lagging at this point. Therefore the transfer function of the perfectly demodulated output velocity, when applied frequency ω is less than, or equal to, the cut-off frequency, can be written as

$$G(j\omega) = \frac{S}{1 + 2Qj\dfrac{\omega}{\omega_v}} \qquad \omega \leq \frac{\omega_v}{2Q} \quad [19]$$

For higher values of ω the amplitude will decrease 6 db per octave but a factor of Q lower than the slope portion of Equation [19]. Therefore

$$G(j\omega) = \frac{S}{2Q^2 j\dfrac{\omega}{\omega_v}} \qquad \omega > \frac{\omega_v}{2Q} \quad [20]$$

These two transfer functions will approximate the output response over the entire range of applied frequencies. Contrary to what might be expected from a second-order system, it is seen that the maximum phase lag is only 90 deg.

FIGURE OF MERIT

An obvious figure of merit for a device of this sort is the product of sensitivity and bandwidth. Therefore, from Equations [9] and [18]

$$S\text{BW} = \frac{\Delta R\omega_v}{R}\left[\frac{I_M}{I_0}\right] \quad [21]$$

in units of velocity output per applied velocity times radians per second.

The sensitivity-bandwidth factor is also a general design equation for possible subminiaturization. As the physical dimensions of a given configuration are scaled down, the inertia ratio will remain constant, the product of $\Delta R\omega_r$ is nearly constant, and decreasing R will give a net increase to the result. Scaling down the output transducer will, of course, decrease the total electrical sensitivity but will not affect performance through loss in resolution.

Steady-State Acceleration Effect

In developing the torque of Equation [4], the applied velocity Ω was assumed to have no time derivative. Later, however, Ω was assumed to be a function of ωt. With this more general function, the torque due to the radial acceleration force term of Equation [28] of the Appendix must be taken into account. Thus the acceleration torque can be written

$$T_a = rF_{0_a} = r^2M\dot{\Omega}(\omega t) \quad [22]$$

Substituting from Equation [2] and retaining only the carrier-frequency term

$$T_a = 2MR\ \Delta R\ \dot{\Omega}(\omega t)\sin\omega_r t \quad [23]$$

Comparing this torque with the torque of Equation [4], it is seen that there is a 90-deg relative phase shift of the carrier. In theory, therefore, the output-velocity component resulting from this acceleration can be removed completely when the output signal is demodulated.

Transient Analysis

If the applied angular velocity consists of a unit step function $u(t)$ the forcing function of the torsion system can be obtained from Equation [28] of the Appendix. With the velocity applied at $t = 0$, the integrodifferential equation for the output velocity V can be written

$$I_0\frac{dV}{dt} + CV + k\int V\,dt = I_0u'(t) + \psi u(t)\cos\omega_r t \quad [24]$$

where $u'(t)$ is the unit impulse and ψ is the coefficient of $\Omega\cos\omega_r t$ in Equation [4]. Neglecting the second term on the right of Equation [24] which would produce the steady-state solution of Equation [9] as well as transient terms, the approximate solution becomes

$$V_T = e^{-\frac{\omega_r}{2Q}t}\cos\left(\omega_r\sqrt{1-\frac{1}{4Q^2}}\right)t \quad [25]$$

While the damped natural frequency of this velocity is different from the undamped natural frequency ω_r the practical physical system would be expected to have even a greater possible difference in value. Therefore the demodulation of this transient component will give the exponential term times a cosine wave of this low difference frequency. The peak amplitude of the transient overshoot is seen to be unity as compared to the steady-state output of Equation [9] which is less than unity. An overshoot of over 100 per cent would thus result. Two methods of preventing this undesirable effect would be either the lowering of the torsion system Q to the critical or overdamped response or by supporting the system housing on mounts with a low transmissibility to applied frequencies near torsion resonance. The latter is a practical method since sensitivity is not reduced and the steady-state output response is virtually unchanged.

Conclusion

The vibratory rate gyro measures angular velocity through the novel application of the characteristics of a vibrating mass. The underlying principles have been fully explored and found to be basically sound and reasonably well suited to practical implementation. The reduction to practice of these unusual principles has been able to draw little on any prior art and, consequently, progress has not been rapid. However, the present state of the instrument's development is encouraging and indicative of a promising future.

Appendix

In vector notation, the velocity of a particle m about a point O moving along a path C in the plane of the paper as shown in Fig. 3 and at a distance r from O can be described in terms of its radial

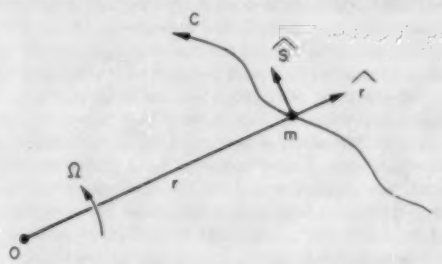


Fig. 3 Radial and Transverse Components of Motion

and transverse resolutes. With unit vectors $\hat{r}$ and $\hat{s}$ in the directions shown, the velocity is

$$\bar{v} = \dot{r}\hat{r} + r\Omega\hat{s} \quad [26]$$

where Ω is the angular velocity. The acceleration is then

$$\bar{a} = \dot{\bar{v}} = (\ddot{r} - r\Omega^2)\hat{r} + (r\dot{\Omega} + 2\dot{\Omega}\dot{r})\hat{s} \quad [27]$$

showing the radial, centripetal, angular, and Coriolis acceleration terms in that order. The reaction force capable of producing a torque about any axis through O is simply

$$\bar{F}_0 = -m(r\dot{\Omega} + 2\Omega\dot{r})\hat{s} \quad [28]$$

or if Ω is constant

$$\bar{F}_0 = -2m\Omega\dot{r}\hat{s} \quad [29]$$

which is due to only the Coriolis acceleration.

Design Features, Operating Experience, and Performance of 100,000-KW Reheat Installation at Potomac River Station

By LEWIS W. CADWALLADER[1] AND HERBERT S. FREDERICK[2]

It is the purpose of this paper to discuss the outstanding features, operating experience, manpower requirements, and performance record of a generating unit with a design capability of 100,000 kw and with steam conditions of 1800 psig pressure, 1050 F primary steam temperature, and 1000 F reheat steam temperature. This design capability of 100,000 kw is somewhat less than the national average of 124,000 kw for all reheat machines purchased in the last 6 years. The initial steam pressure and temperature have been surpassed by other installations, and condenser circulating-water temperature is relatively high, with maximum temperatures of 88 F and a 4-month average of 81 F for June to September, inclusive. Notwithstanding these basic limitations, the over-all thermal efficiency is averaging 37 per cent which is near optimum for conditions below the supercritical. The unit under discussion is the No. 3 unit in a station housing also two 90,000-kw nonreheat machines placed in operation in October, 1949, and June, 1950. The new No. 3 unit was placed in preliminary operation June 30, 1954, and became firm commercial capacity 2 weeks later on July 14, 1954. Unit No. 4, which is a duplicate of No. 3, is under construction and is scheduled for service early in 1956. There is a single boiler per turbine, and fuel used is pulverized coal. Light fuel oil is used for ignition.

Introduction

THE design criteria established for the Potomac River Generating Station called for compliance with the varying and exacting requirements of such authorities as the National Capital Parks Division of the U. S. Department of the Interior, the City of Alexandria, Va., and the Civil Aeronautics Administration. The obedience to the dictates of these bodies, even when overlapping and conflicting, has been described elsewhere.[3-5]

The reasons for these extraordinary stipulations are clear when we consider that the Potomac River Generating Station of Potomac Electric Power Company is located on the Potomac River within 5 miles of the Capitol of the United States, 1.55 miles from the nearest part of the Washington National Airport, and approximately 250 yd from several groups of modern apartment buildings. This location made it mandatory that the station building harmonize with the general architectural pattern of the area. This was accomplished by the use of red exterior face brick trimmed with gray limestone. The chimneys were extended only 161 ft above the ground but equipped with nozzles which at full load discharge the gases at 60 mph.

Each boiler is provided with its own chimney. Station noise carry-over to the apartment buildings is controlled and minimized by the use of glass-block windows on that side of the building, by using slow-speed exhaust fans, and by installing silencers on the usual sources of station noises such as motor-driven vacuum pumps, steam flash tanks, and ash-conveyer exhausters. Fly-ash emission from the chimneys is controlled by a series mechanical-electrostatic dust collector with a combined collection efficiency of 99 per cent. The rotary railroad-car dumper is totally enclosed and equipped with a dust-collecting system to permit dustless unloading of coal cars. The Bradford breaker house, coal bunkers in station and all intermediate coal-conveyer transfer points also are served by dust-collecting systems. The surface of the 150 000-ton coal-storage pile is sprayed periodically with heavy fuel oil to prevent dusting. The station and the neighboring area are shown in Fig. 1.

FIG. 1 POTOMAC RIVER STATION

[1] Mechanical Engineer, Potomac Electric Power Company, Washington, D. C. Mem. ASME.

[2] Mechanical Engineer, Stone & Webster Engineering Corporation, New York, N. Y. Mem. ASME.

[3] "More Power for Our Nation's Capital," by L. W. Cadwallader and W. J. Lank, *Combustion*, March, 1950.

[4] "Potomac Electric Power Company's New Potomac River Station," by L. W. Cadwallader, *Southern Power & Industry*, February, 1951.

[5] "Cost, Application, and Performance of Series Dust Collector Installations on Large Pulverized Coal Boilers," by L. W. Cadwallader, paper presented at the 1952 Annual Convention of the Air Pollution and Smoke Prevention Association of America.

Contributed by the Power Division and presented at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–21, 1955, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 20, 1955. Paper No. 55—S-21.

Main Generators

Each of the main generators is directly connected to two transformers to step up the voltage from 13,800 to 69,000 volts for transmission to the load. Two 60,000-kva transformers are provided for each of the 90,000-kw generators and two 90,000-kva transformers for each of the 100,000-kw generators. The 13,800-volt generator leads to the transformer terminals are of the isolated phase construction. The normal source of supply for the 2300-volt station auxiliaries is from transformers directly connected to the 13,800-volt leads of each main generator. There is a reserve source of supply to these auxiliaries through transformers connected to the 69,000-volt bus. The outdoor switchyard is adjacent to the turbine room and may be seen in Fig. 1.

Nos. 1 and 2 units were provided originally with steam evaporators for furnishing distilled water make-up. A mixed-bed demineralizer was included in the job for Nos. 3 and 4 units but is not yet in service. This demineralizer will have sufficient capacity to provide all make-up for an ultimate station of five units.

In conjunction with the basic economic studies for this new No. 3 unit, careful consideration was given to selecting steam conditions and a cycle arrangement which were progressive enough to warrant duplication in a fourth and fifth unit at Potomac River Station and at the same time conservative enough to give some assurance of reliable service and low operating and maintenance costs. A substantial saving in engineering costs can be realized by a realistic policy of duplication within the framework of good economics.

Unit No. 3

Based on the foregoing considerations, the decision was made to proceed with a unit having a design capability of 100,000 kw, 1800 psig pressure, 1050 F primary steam temperature, 1000 F reheat steam temperature, seven stages of extraction feedwater heating with a maximum feedwater temperature of approximately 450 F entering the boiler economizer. A cross section of the station is shown in Fig. 2.

The fundamental flow diagram of unit No. 3 is shown in Fig. 3. The first points of regenerative feedwater heating are the 22nd and 21st stage heaters located in the steam-dome space of the condenser. The two open heaters provide for deaeration and boiler-feed storage. Final feedwater heating is accomplished in two half-capacity heaters connected in parallel and supplied with steam from the cold reheat lines. Two heaters in parallel were selected to avoid a drop of 84 deg F (453–369) in final feedwater temperature which would occur if a single full-capacity heater were to be lost in this service.

Turbine Generator. The turbine generator selected was a General Electric 3600-rpm, tandem-compound, double-flow reheat machine of the so-called Dunkirk design as shown in Fig. 4. The machine was specified to have a guaranteed maximum capability of 100,000 kw at 2.0 in. Hg abs back pressure. To comply with this type specification the manufacturer designed the machine with a theoretical margin of 5 per cent. The hydrogen-cooled generator was specified to have a maximum guaranteed capability of 100,000 kw at 0.85 power factor, 13,800 volts, and with 95 F river water supplied to the hydrogen coolers. Excitation is by a 300-kw, 1200-rpm, General Electric high-inertia-type motor-generator set with 250-volt main and pilot exciters.

Boiler The boiler selected was a Combustion Engineering controlled-circulation unit rated at 725,000 lb of steam per hr. at 1875 psig superheater-outlet pressure, 1050 F primary steam temperature, and 1000 F reheat temperature. A boiler cross section is shown in Fig. 5. The exit-gas temperature was speci-

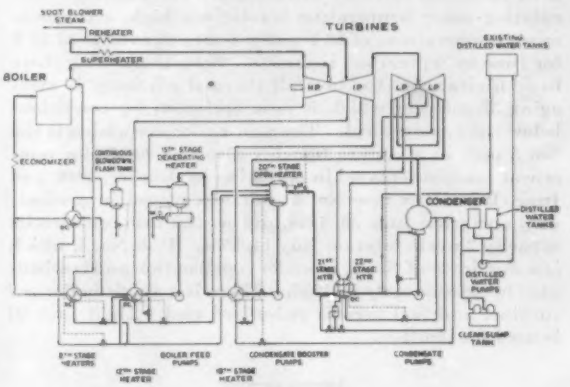


Fig. 3 Fundamental Flow Diagram—No. 3 Unit

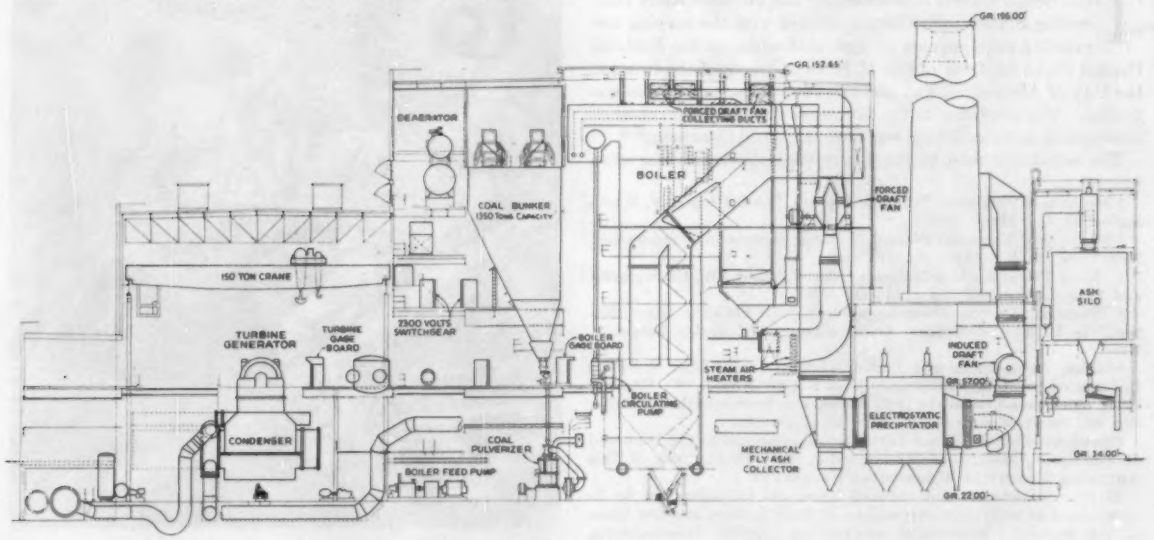


Fig. 2 Cross Section

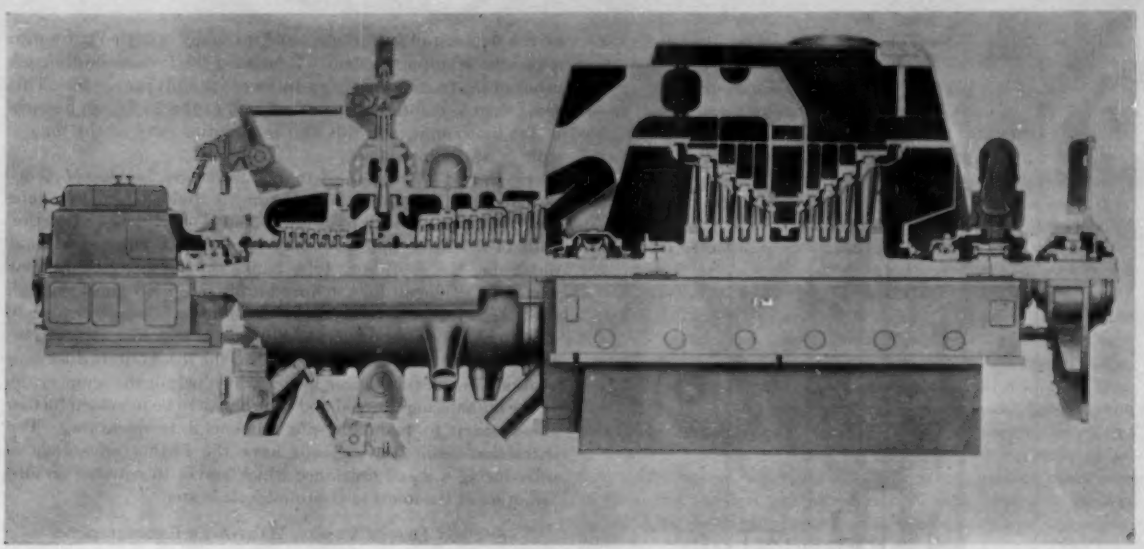


FIG. 4 CROSS SECTION—NO. 3 TURBINE

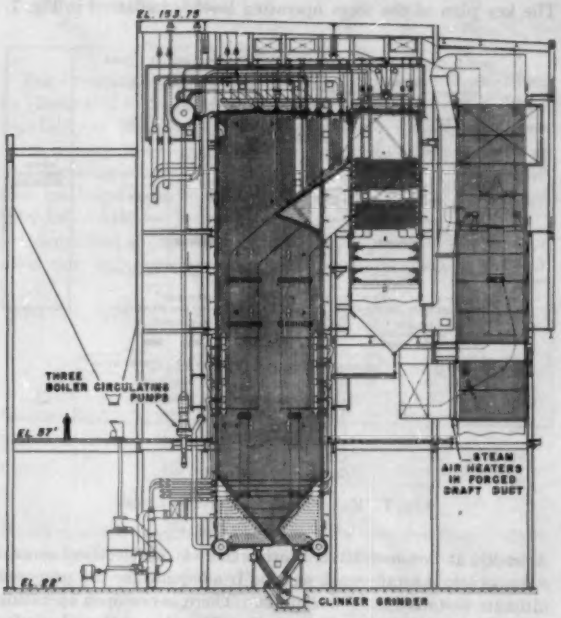


FIG. 5 CROSS SECTION—NO. 3 BOILER

fied at 250 F with 100 F air entering the tubular air heater, and the boiler setting was specified for zero air infiltration. The intent of the boiler specification was to cover a steam-generating unit having the maximum efficiency attainable within the limit imposed by the gas-exit temperature of 250 F. There has been reluctance on the part of many engineers to purchase a steam-generating unit for this low gas temperature because of the possible corrosion and plugging of the air preheater. We felt that with the good-quality coal available to this plant it would be a calculated risk well worth taking. As a result, a boiler unit was developed by Combustion Engineering with a guaranteed efficiency of 91.2 per cent when burning eastern bituminous coal with a heat content of 13,000 Btu per lb as fired.

Mr. W. H. Armacost presented a paper[6] at the 1953 Annual Meeting of the Society which described controlled-circulation boilers then in operation. Results of steam-purity tests and observed drum-metal temperatures during start-up of our No. 3 boiler were reported at the 1954 Annual Meeting of the Society by Messrs. Powell and Grabowski.[7] Only features differing from earlier units or of particular interest pertaining to labor saving and performance will be covered in this discussion.

Boiler Circulating Pumps. There are three motor-driven Ingersoll-Rand vertical circulating pumps per boiler, any two of which are capable of circulating adequate water for full steam output. Each pump is rated at 4250 gpm at a pump discharge pressure of 1985 psig with a pump suction pressure of 1945 psig, and is driven by a 150-hp, 1750-rpm motor. A cross section of one of the boiler circulating pumps is shown in Fig. 6

These pumps are of the vertical type, with single-suction impeller, double-volute casing to eliminate radial thrust. Suction and discharge nozzles are arranged for welding to the boiler circulation piping which supports the pumps. Casing material is cast-carbon steel. The casing cover is forged carbon steel drilled for necessary connections. Pressure bolting is heat-treated alloy steel; this bolting is accessible for easy removal with an impact wrench. Removal of pressure bolting nuts allows the complete pumping element to be removed without disturbing the basic pump alignment. The shaft is entirely supported by external bearings—ball-type radial bearings and angular contact roller-type thrust bearings. Impeller material is chrome-moly steel with renewable impeller and casing wearing rings. The sealing arrangement consists essentially of a throttle bushing and sleeve between the injection-water chamber and the back of the impeller. The function of this seal is to prevent high-temperature boiler water from entering the breakdown device. Provision is made to install a full pressure mechanical seal without injection water if a

[6] "The Controlled-Circulation Boiler," by W. H. Armacost, Trans. ASME, vol. 76, 1954, pp. 715–725.

[7] "Drum Internals and High-Pressure Boiler Design," by E. M. Powell and H. A. Grabowski, Combustion Engineering, Inc. ASME Paper No. 54—A-242.

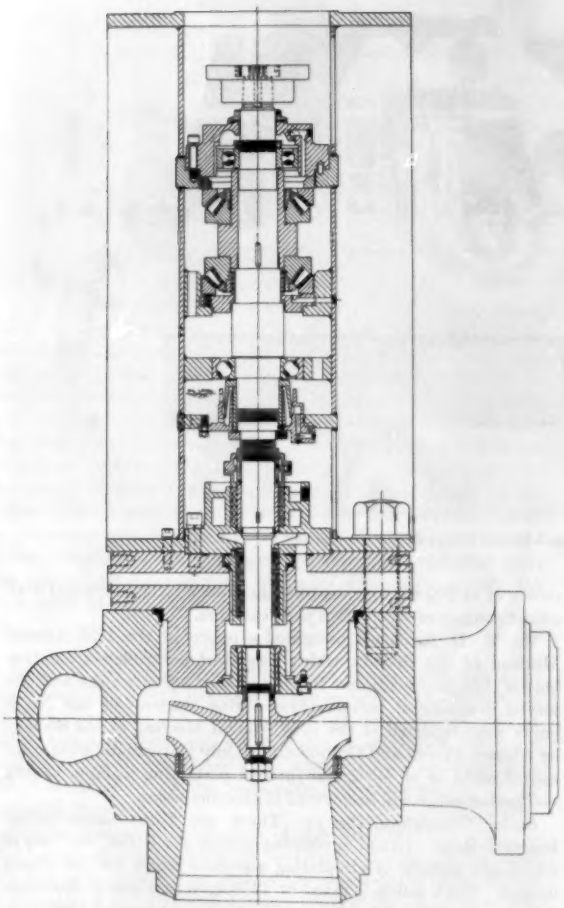


FIG. 6 CROSS SECTION—BOILER CIRCULATING PUMP

satisfactory mechanical seal is developed. A floating seal pressure-breakdown device is provided between the injection water chamber and the stuffing box.

In the design stage, it was expected that the high-pressure injection-water requirement would be 25 gpm per pump or 75 gpm for three pumps. In order to determine this requirement accurately during normal boiler operation, the boiler feedwater flow to the common header furnishing injection water to all three pumps is recorded by a flowmeter. The leakoff to the 15th stage deaerating feedwater heater is similarly metered and recorded. Also a differential-pressure control valve was installed to maintain a prescribed differential pressure across the pump seal regardless of injection water-supply pressure variations or of pump suction pressure. We are operating with a constant differential pressure of 40 psi across the pump seal, under which condition the total injection water supply for three pumps is 50 gpm and the total leakoff for three pumps is 37 gpm. The difference of 13 gpm or 6500 lb per hr represents the total quantity of feedwater flowing through the three pump seals into the boiler.

Boiler Furnace. The cross section shown in Fig. 5 indicates that the furnace has a dry bottom with shaded ashpit. This design of furnace bottom has resulted in a substantially cooler ashpit than the conventional symmetrical design. The furnace-bottom ash is removed in a completely dry state through motor-driven double-roll ash grinders and manually operated gates into a vacuum transport system. Removing the furnace-bottom ash is one of the routine operating duties of the shift personnel. This same system empties automatically all of the 28 fly-ash hoppers of the boiler unit, and this also is a routine duty of the regular shift operating personnel.

Air Heater. The vertical tubular air heater, because of its design for 250 F gas-exit temperature, is divided into two tube banks—one bank with 56-ft tubes and a short cold-end section with 14-ft tubes. The 14-ft tube bank can be renewed without disturbing the 56-ft-long tube bank, should this ever become necessary. The air for the forced-draft fans is obtained through duct work from above the boiler in order to furnish warm air to the tubular air heater. In addition to this precaution, extraction-steam heating coils are located in the forced-draft duct connections to the tubular air heater which permit the temperature of the air entering the cold-end tube banks to be increased further if necessary to maintain safe tube-metal temperatures. The extraction-steam heating coils have the further advantage of introducing a small resistance which serves to equalize air distribution at the inlets to the tubular air heater.

STATION DESIGN VERSUS MANPOWER REQUIREMENTS

The station layout does not embrace the central control-room idea to the extent so common in new generating stations of today. The key plan of the main operating level is indicated in Fig. 7.

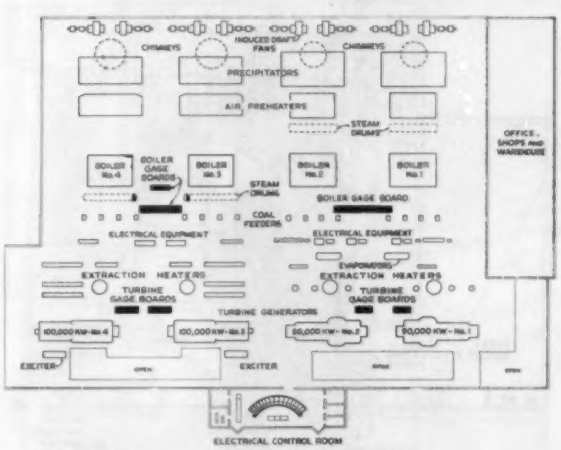


FIG. 7 KEY PLAN—OPERATING LEVEL

Actually, at Potomac River Station there is one electrical control room which is totally enclosed and is adequate for the proposed ultimate installation of five units. There is one open operating station in the boiler room for boilers Nos. 1 and 2, and one for boilers Nos. 3 and 4. As shown in Fig. 8, there is one open operating station in the basement for Nos. 1 and 2 auxiliaries and one for Nos. 3 and 4 auxiliaries. There are no regularly assigned turbine operators. Starting and stopping of machines are performed by a steam operator from the boiler room under supervision of the watch engineer or the assistant watch engineer. This means that for the four units at Potomac River Station, three of which are in service and one of which is under construction, there will be a total of five principal operating stations—two in the boiler room, two in the basement, and one in the electrical control room. The operating complement of this four-unit station, including all shift operating supervision, electrical operators, and steam operators will be as follows:

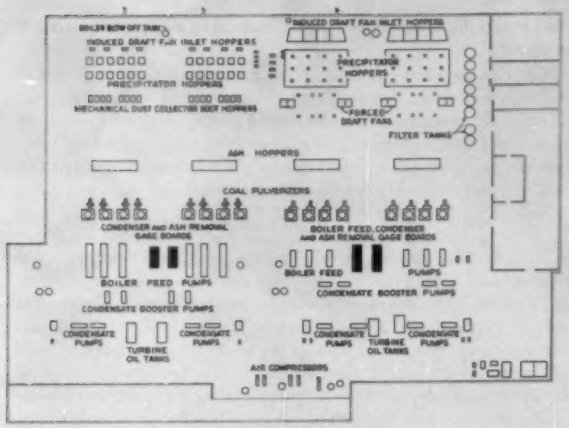


FIG. 8 KEY PLAN—BASEMENT LEVEL

	Each 8-hr shift	Total
Watch engineers	1	5
Assistant watch engineers	1	5
Steam operators	8	40
Electrical operators	2	9
	12	59

The foregoing is the total operating-shift personnel which will be chargeable to the operation of these four units with a design capability of 380,000 kw. There are no other equipment operators, roving attendants, relief operators, operators in training, or maintenance men who charge part time to plant operation. The total operating complement of 59 provides 0.155 man per 1000 kw or 6440 kw per man.

Comparison of operating manpower requirements with those of other new pulverized-coal-burning stations is given in Table 1.

TABLE 1 OPERATING MANPOWER REQUIREMENTS—SHIFT OPERATING PERSONNEL ONLY

Plant	Plant capability, megawatts	Number of turbines	Number of boilers	Type of fuel	Shift personnel actual payroll
Potomac River	380	4	4	Pulverized coal	59
Plant A	640	4	4	Pulverized coal	64
Plant B	660	4	4	Pulverized coal	56

Plants A and B have larger-capacity units operated from centralized control rooms. It is our considered opinion that the operating complement of 59 for Potomac River Station could man a station of similar design containing four 200,000-kw machines, in which case the men per 1000 kw would become 0.074 and the kilowatts per man would become 13,560.

We believe that the operating organization and arrangement of control stations at the Potomac River Plant compare favorably with central control-room stations now in operation, and have the added advantages of lower capital investment for instrumentation, lower instrument maintenance costs, and distribution of the operating personnel to better advantage around the station for both normal and emergency operation.

OPERATING EXPERIENCE AND PERFORMANCE

In accordance with our past practice for new boilers, the No. 3 boiler unit was acid-cleaned. One of the Ingersoll-Rand vertical-boiler circulating pumps was used throughout the acid-cleaning procedure without any special parts or undue attention.

All main steam and reheat steam piping was thoroughly blown out to atmosphere with live steam from the boiler before any steam was admitted into the turbine. As an additional safeguard, temporary fine-mesh strainers were installed around the permanent steam strainers in the main steam stop valves and in the reheat-steam intercept valves for a period of 10 days during preliminary operation. They were removed at the expiration of this 10-day period because of our concern over previous reports from other utilities indicating disintegration of the fine-mesh wire after very short periods of operation. However, we found our fine-mesh strainers in excellent condition with every indication of being suitable for a much longer period of operation than 10 days. The four fine-mesh strainers had accumulated approximately 0.50 oz of welding beads and fine metal chips, some of which could have caused damage to the turbine blading. We believe that the omission of backing rings contributed to the cleanliness of the piping systems of the main steam, reheat-steam, and boiler feed-pump discharge lines. On those systems all joints were welded under a procedure in which backing rings were omitted and the first two passes in each pipe joint were applied by an inert gas-shielded arc-welding technique using argon gas; from that point the standard electric arc-welding process was used to complete the joint.

No unusual difficulties were experienced in placing in operation this our first reheat installation. The operating force, which had previously operated only nonreheat units, managed the new reheat job very well and with little or no difficulty.

Outages. There have been two outages since this unit was made firm commercial capacity on July 14, 1954. The first was a scheduled 5-day outage (September 2–7, 1954) to replace a faulty casing on one of the boiler circulating pumps. The fault in this pump casing was apparently a foundry crack which was not discovered during foundry and shop examination of the casting. The crack was found during a routine examination of the pump and not because of any evidence of distress. Although the pump is equipped with shutoff valves to permit normal maintenance with the boiler in service, it was deemed inadvisable to renew the pump casing with the boiler under pressure because of the fact that the discharge angle-type stop-check valve is welded directly to the pump-discharge nozzle.

The second interruption was a forced outage on December 7, 1954, caused by a burn-up of the positive collector rings on the main generator. This trouble developed between 4:00 p.m. when the rings were inspected and no visible evidence of distress was noted, and 4:40 p.m. when the positive rings were found white-hot and had already partially melted the carbon brushes and some of the brush holders. The unit was carrying 100,000 kw at the time. It was shut down at 4:46 p.m. and remained out of service until December 12, 1954, for installation of new rings and brush rigging. We understand that there were other similar failures around the country in 1954—all on the positive rings of 3600-rpm generators. In each case the construction consisted of two positive rings with brushes operating in parallel. The negative rings are similarly constructed but have not experienced this difficulty. The cause of the failures of the positive rings is unknown at this time.

During each of the outages, the cold end of the tubular air heater was inspected carefully to observe the effect of the 250-F gas-exit temperature. The first inspection, made after approximately 2 months of operation with an average air-inlet temperature of 110 F, revealed a somewhat heavy coating of fine ash but no plugging or corrosion. As a result of this inspection and as an additional safeguard, the air temperature to the tubular air heater was increased to an average of approximately 125 F by use of the

extraction-steam heating coils. The second inspection, made approximately 3 months after the first inspection, indicated the tubular air heater to be in first-class condition with only a thin coating of fine ash.

Station Heat Rates. The net station heat rates for unit No. 3, based on design calculations, are shown in Fig. 9. The actual operating results for the period from July 14, 1954, to March 31, 1955, are given in Table 1-A. During this period the average operating turbine load was 98,000 kw and the average over-all net heat rate was 9190 Btu per kwhr.

TABLE 1-A POTOMAC RIVER STATION OPERATING REPORT—NO. 3 UNIT

(From 4:00 p.m., July 14, 1954, to midnight, March 31, 1955)

DESCRIPTION		
AVERAGE TURBINE LOAD		98,000 KW
STATION SERVICE-% OF GROSS GENERATION		5.42 %
MAIN STEAM PRESSURE AND TEMPERATURE	1,876 PSIG	1,032 F
HOT REHEAT PRESSURE AND TEMPERATURE	384 PSIG	988 F
FEEDWATER TEMPERATURE		455 F
CONDENSER BACK PRESSURE		1.11 In Hg.
CIRCULATING WATER TEMPERATURE		56 F
MAKE-UP-% OF BOILER FEED		0.288 %
EXIT FLUE GAS TEMPERATURE		256 F
BOILER STEAMING EFFICIENCY		91.2 %
BOILER USE FACTOR		96.3 %
POUNDS OF COAL PER KILOWATT HOUR NET		0.649 Lb
BTU PER KILOWATT HOUR GROSS		8,692 Btu
BTU PER KILOWATT HOUR NET		9,190 Btu
THERMAL EFFICIENCY		37.14 %

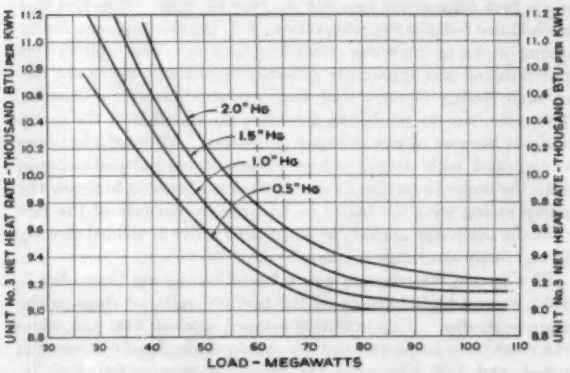


FIG. 9 NET STATION HEAT RATES

The most efficient steam station reported by the Federal Power Commission for the year 1953 had a net heat rate of 9170 Btu per kwhr. The units in that station, however, are 175,000-kw cross-compound units with steam conditions of 2000 psig, 1050 F/1050 F, which is inherently a more efficient design than our 100,000-kw tandem-compound unit designed for 1800 psig, 1050 F/1000 F.

Controlled-Circulation Boiler Performance. The controlled-circulation boiler unit is worthy of particular mention in that its design factors and performance have been almost exactly as predicted. The boiler steaming efficiency for the 8½-month period covered by Table 1-A was 91.2 per cent. The coal burned during this period had an average heat content of 14,162 Btu per lb as fired and an average sulphur content of 0.70 per cent. Steam temperatures have been controlled by burner tilt and an occasional use of steam soot blowers. No spray water, either for main-steam or reheat-steam temperature control, has been used except for the purpose of calibrating the spray water-control systems.

We wish to point out that all integrating watthour meters are carefully calibrated, automatic samplers are used for coal samples, all car coal is weighed on accurate track scales, and weekly analysis of coal is performed in our own laboratory with a monthly verification by four other laboratories. Therefore we believe the performance reported for this unit to be true and representative.

CONCLUSIONS

Each advance in steam-cycle arrangement and in pressure and temperature entail design problems and present potential operating difficulties and the possibility of increased maintenance costs. It is appropriate, therefore, that, within the limits of realistic economic returns, maximum performance be exacted at each pressure classification. Particularly is this true in the range of pressures and temperatures common today, for there are diminishing returns as further advances are made in steam conditions. Furthermore, no matter how progressive and efficient the design may be, a generating unit is not earning any money when shut down for repairs or revisions.

The record of the unit under discussion is evidence of what can be done if careful attention is given to all details in selection of equipment, station design, and operation. The demonstrated performance, at steam conditions that are now generally accepted as somewhat conservative for utility stations, may be taken as an attainable datum for evaluating the advisability of assuming the costs and risks inherent in more advanced designs.

Discussion

W. W. BROWN[8] AND R. M. VAN DUZER.[9] The authors have presented an interesting paper showing how relatively high plant thermal efficiency can be gained "without assuming costs and risks inherent in more advanced designs." It seems possible, however, that higher-design throttle pressure using the same primary and reheat temperatures indicated in the paper would yield worth-while thermal economy. Pressures up to 2350 psi have been used in several installations, without impairing reliability.

Since high boiler efficiency, 91.2 per cent, contributes largely to the excellent net heat rate for the unit, we should like further explanation on this subject. We know it is difficult to compare boiler efficiency in different areas of the country mainly because of solid fuel quality and its effect on efficiency assuming common stack temperatures. For example, we have now on order a large steam generator designed for 200 F stack temperature. With this temperature and using 80 F air ambient and lower-quality coal, we expect an efficiency of 90.2 per cent or 1 per cent lower than the boiler in question. We mention this to illustrate the important effect of coal quality and air ambient.

It might be helpful, therefore, if the authors would explain the factors contributing to the boiler manufacturer's efficiency guarantee, 91.2 per cent. One of these, of course, is the lower than average stack temperature of 250 F which will result in lower dry-gas loss and improved efficiency, but how about the others? In this case, was a higher ambient used than the 80 F generally accepted? In our most modern units we specify coal with 8 to 9 per cent moisture and 5 per cent hydrogen when requesting boiler-efficiency guarantees. How does this compare with the percentages used in the authors' case?

[8] Chief Mechanical Engineer, Mechanical Engineering Division, The Detroit Edison Company, Detroit, Mich. Mem. ASME.

[9] General Superintendent of Production, The Detroit Edison Company, Detroit, Mich. Mem. ASME.

The author states that the boiler setting was specified for zero air infiltration. It is assumed that the unit is a balanced-draft unit. The high boiler efficiency guaranteed and attained in actual operation indicates that the infiltration is a negligible amount. One of the major economies claimed for pressurized furnaces is the savings attributed to low-infiltration losses. We have been convinced that this saving can be realized on a balanced-draft unit by careful attention to casing design. Results at the Potomac River Station substantiate this conviction.

We find considerable interest in the manpower figures stated in the paper. The men per kw compare very favorably with our St. Clair Station figures of 0.102 men per Mw. Four 156,000-kw units are now in operation, two units being operated from each of two control rooms. Considerations other than manpower, however, may lead to selection of centralized control, among these are improved operating conditions resulting from air-conditioned enclosed rooms, cleaner instruments with better accuracy and less maintenance. We feel, also, that better protection for operating in emergencies is important, particularly during a fire since the control-room air should be taken from outdoors and the room pressurized sufficiently to keep out smoke. It is our opinion that these things are worth the cost of the room enclosure. As a matter of interest the manpower ratio per Mw for our River Rouge Plant, now under construction, will be 0.080 based on three units totaling 820 Mw. This illustrates what size effect does to the ratio.

We believe, also, that conservative planning for units of 200 Mw and upward calls for additional manpower coverage for each unit as insurance against outage considering the large block of capacity involved.

We are most interested to learn that the omission of backing rings on the main steam, reheat, and boiler-feed discharge lines, and the use of inert arc welding resulted in relatively clean piping systems so that a negligible amount of welding beads and fine welding chips was accumulated in the strainer. It would be helpful to know whether radiographic examination was used and, if so, at what stage or stages of completion of the welds.

N. R. Deming.[10] It was pointed out that for a period of 8½ months the average over-all net heat rate was 9190 Btu/kwhr, a very desirable value. However, it would seem that when comparing the heat rate with those obtained in other steam stations, a correction should be made, at least for the 4.2 per cent excess throttle pressure over the design value. It is not always possible to control superheater and reheater temperatures to the desired values and, of course, any station heat rate is entitled to be benefited by the gain from having less than the design value of turbine-exhaust pressure.

Station heat rates are the result obtained from two kinds of measurements—fuel burned or heat added in the combustion process and net generation or station output. Obviously the accuracy of the heat-rate figure depends directly on the accuracy of each of these two measurements.

Without making specific reference to the Potomac River Station the writer wishes to question the reliability of load measurements taken from integrating watthour meters, even though they may be considered to be calibrated carefully. Irregular discrepancies between values from a calibrated watthour meter and the three-wattmeter method of load measurement of up to 1 per cent during acceptance economy tests of the writer's experience prompt this question. While it may be that discrepancies of this order of magnitude occur infrequently, it would be interesting to know what accuracy is expected by the electric utilities in heat rates computed from data collected over a long period of time.

E. H. Krieg[11] and W. F. Ryan.[12] The generating plants of Potomac Electric Power Company have a long tradition of sound engineering design and expert operation. This latest installation is in strict conformity with that tradition, and illustrates what gratifying results can be obtained with what the authors call a "conservative" design by meticulous attention to thermodynamic and mechanical details when combined with carefully planned and skillfully executed operating procedures. This plant also maintains the Pepco tradition for fine architecture, cleanliness, and quietness so essential in the nation's capitol, and also high output per operator. It is impossible to make direct or conclusive comparisons on operating personnel, but we doubt that any utility has been more efficient in respect to manpower, and the company's excellence in this respect dates from the earliest installation at Buzzard Point.

The plant is noteworthy for some features that are not in other plants or, if so, have not come to our attention. Unusual advantage has been taken of the potentialities of the low end of the heat cycle which has compensated to some extent for the "conservative" throttle condition. This has been a factor in achieving a performance of 9190 Btu per net kwhr output, in spite of a condenser back pressure of 1.36 in. Hg. This record was made possible by a 91.3 per cent boiler efficiency and a stack temperature of 256 F. The plant has enjoyed an unusually high availability with no outages chargeable to the cycle.

There is one particular in which Potomac River Station has occupied a unique position, in our experience, since the first unit went on the line, and that is its outstanding reputation for being "a good neighbor." Fig. 1 of the paper shows a group of apartment houses situated almost up to the plant property line. At the upper left-hand corner are the lights indicating the approach to the Washington airport. Another apartment house erected a short distance from that shown in the photograph, was built with white brick on the south side to help keep apartments cool in the summer, and after several years the white brick is still clean.

As the plant is close to the approach zone of the Washington Airport, the stacks had to be kept very low, about 160 ft above ground level. To make this possible, a combination of mechanical and electrical electrostatic dust collectors was used which, as the authors report, have attained the remarkable collection efficiency of 99 per cent. Those who have flown over the plant when going to Washington by plane can verify the statement that on such occasions it would seem that one could look to the bottom of the stack even when the plant is carrying full load.

In these days, when many utilities are having trouble with zoning boards, the Potomac River Station presents an outstanding example of the fact that a generating station, designed for its particular environment, can be a good neighbor and not a nuisance.

R. C. Roe[13] and M. A. Mayers.[14] The authors are to be congratulated on their description of the new addition to the Potomac River Generating Station. It is especially noteworthy that the performance of the plant, for its first 8½ months of operation, is so close to the anticipated performance shown in Fig. 9 of the paper. This fact emphasizes the ability of design

[10] Design Engineer, Steam Engineering Department, Westinghouse Electric Corporation, Philadelphia, Pa. Assoc. Mem. ASME.

[11] Consulting Engineer, Stone & Webster Engineering Corporation, Boston, Mass. Fellow ASME.

[12] Senior Consulting Engineer, Stone & Webster Engineering Corporation. Honorary Mem. ASME.

[13] President, Burns & Roe, Inc., New York, N. Y. Fellow ASME.

[14] Chief Mechanical Engineer, Burns & Roe, Inc. Mem. ASME.

engineers and equipment manufacturers to accomplish the objectives of their predicted design in the American power industry. This, in turn, permits power companies to design with assurance for the thermal economy appropriate to the state of development of their operating systems and thus to assure themselves of a maximum dollar return on their investment.

A noteworthy feature of the heat balance shown in Fig. 3 of of paper is the use of a direct-contact heater at the 20th stage extraction point in addition to the usual direct-contact deaerator heater at the suction of the boiler feed pump. This practice approaches the economy inherent in the direct-contact heater system which the writers' company has used on many of its high-efficiency stations where heaters are of the direct-contact type for all stages below the boiler-feed-pump suction. The three pumps operating on the condensate cycle would seem to have some disadvantages compared with the two pumps in the standard system of the writers' company. With the inherent present-day high reliability and availability of properly designed multisection condensate-heater pumps it would seem that full advantage could be taken of a full direct-contact heater system if it were desired.

Having gone so far, it would be of interest if the authors would explain why they did not use open heaters at the 21st and 22nd stage extraction points. We wonder also, in view of the seven stages of feedwater heating, why the temperature of the feedwater heater was not taken above 450 F. Our own analyses indicate that an additional 30 to 50 deg of feedwater heating probably would pay off in an area of high fuel cost.

The low boiler flue-gas-outlet temperature of 256 F is evidently made possible by the high quality of coal utilized in this station. This permits going to an extraordinarily high efficiency which could probably be duplicated in few other plants. Are enamel tubes being used in the air preheater to help eliminate corrosion? It is interesting to note that no evidence of plugging or corrosion of the cold end of the air preheater was found after 3 months of operation. Can the authors give the result of any more recent observations?

As the authors indicate, the omission of a central control room is somewhat unusual in a new steam power plant. While the stated manpower figures indicate economical manning in spite of this omission, we wonder whether some difficulties may not have appeared during emergency or transient conditions of operation.

We wonder also whether the authors could quote, at this time, the cost per kilowatthour of name-plate capacity of this new addition. In view of the special conditions of cleanliness and quiet that this plant must meet it would be expected that the construction cost would be somewhat high. It would be of interest to see how much increase in cost is necessitated by these special requirements.

Authors' Closure

The authors appreciate the general interest shown in this paper both by those who presented written discussion and others who offered comments or asked questions during the session.

The boiler losses predicted in the boiler contract and indicated by two heat-balance tests in field are given in Table 2.

TABLE 2

Losses, per cent	Boiler contract	Heat-balance tests
Dry gas	3.38	3.50
Water and hydrogen in coa	3.21	3.57
Water in combustion air	.09	.10
Unburned carbon	.30	.27
Radiation	.32	.32
Unaccounted-for losses	1.50	1.50
Total losses	8.80	9.26

The boiler specification and contract were based on 100 F air from the boiler room entering the forced-draft fans. The ultimate analyses of the coal samples obtained during the heat-balance tests indicated average moisture and hydrogen contents of 4.46 per cent and 4.47 per cent, respectively.

Due to the advantageous location of the 21st and 22nd stage closed feedwater heaters in the steam-dome space of the condenser, we do not believe that direct-contact heaters with associated pumping equipment would have been economically feasible in this particular installation.

We considered final feedwater temperatures up to 500 F but concluded that 450 F represented the most economical installation, as feedwater temperatures in excess of 450 F resulted in higher boiler costs which we could not justify.

Carbon-steel tubing is used throughout the tubular air heater, and another inspection continues to indicate a very satisfactory condition. We believe this can be attributed to two things—good quality coal and 125 F air entering the tube banks.

We gave serious consideration to providing a central-control room for the two new reheat units in this station. However, we were not able to satisfy ourselves that this was the economical thing to do. Further, we believe there are advantages to all plant operators being out in the station where they can see and hear. It is our honest opinion that operators distributed as at the Potomac River Station are in a better position to handle emergency conditions than they would be if housed in an enclosure. We have been able to handle all emergencies quite satisfactorily with the operating arrangement described. It may be that an enclosed and properly ventilated central-control room could have an advantage in event of a serious fire. On the other hand, the fire might be discovered and controlled sooner by operators located out in the station.

The cost of this addition, consisting of two 100,000-kw design-capability reheat units, including the switchyard, will be very close to $31,000,000, or $155 per kw.

Visual inspection by a borescope was made of the inside condition of the first bead applied in the pipe welding by the inert-gas-shielded process. Any defects located were repaired by remelting the areas from the outside. A second visual inspection was made following the second bead. The welds were completed by the arc-welding process and radiographs made of the completed welds.

Messrs. Roe and Mayers suggest the possibilities of using open heaters at the 21st and 22nd stage extraction points and 30 to 50 degree higher final-feedwater temperature. Adoption of these changes would have resulted in higher cost and less convenience in operation which we consider not warranted in our case for the expected higher efficiency. A drain cooler in the 22nd stage heater limits the gain available from open heaters and location of the closed heaters in the condenser simplifies station layout and costs less.

The 4.2 per cent excess throttle pressure mentioned by Mr. Deming actually was the pressure at the superheater outlet. A better comparison of design and operating throttle conditions would be the heat contents which are 1510.9 and 1499.0 Btu per lb, respectively, or approximately 0.8 of 1 per cent below design conditions.

An Approach to the Economic Problem of Matching Condenser Surface With Exhaust-Annulus Area

By W. A. WILSON[1] AND L. G. MALOUF[2]

The concept of economic balance as related to power-plant optimization is discussed. Its application to the problem of choosing appropriate exhaust-annulus and condensing surface areas is developed in detail. These areas are related to each other for three circulating-water temperatures and several unit-cost ratios. The exhaust conditions and the rate of change of stagnation enthalpy with investment are presented as functions of these areas.

Introduction

IT IS the purpose of this paper to present a method of analysis whereby the separate tasks of selecting the appropriate condensing surface and the appropriate turbine-exhaust-annulus area for a given power-plant application are reduced to the single problem of selecting a balanced combination of these areas. This simplification is possible because the effects of each of these areas on operating cost and on plant capacity are made manifest by their influences on the same working-fluid property, i.e., the stagnation enthalpy leaving the last turbine-blade row. At the same time, the remaining components of the plant are sensitive only to this enthalpy and not to variations of the combination of condensing and exhaust-annulus areas which establishes that enthalpy. This means that if the stagnation enthalpy is assigned a value, the corresponding economically appropriate condenser-annulus combination is the one which produces that enthalpy for the least capital investment. This combination is balanced for the particular enthalpy. A balanced combination exists for each value of stagnation enthalpy. These combinations are independent of the design of the remainder of the plant.

The analysis which follows develops not only the balanced relationship between the two areas but also the rate of return on marginal investment in these plant elements for each balanced combination. This is expressed in terms of improvement in stagnation enthalpy per marginal dollar invested for each unit of condenser flow. This information is directly applicable to the selection, on economic grounds, of the particular balanced combination appropriate for a given power plant.

Formal Analysis of Condenser and Exhaust-Annulus Balance

When the exhaust-annulus area and condensing surface are balanced, an incremental reduction in the condenser surface will result in an increase of the enthalpy at the state-line end point which is exactly compensated by the reduction in exhaust loss resulting from applying the saving in condenser investment to an increase in exhaust-annulus area. Under this condition the same power will be developed for the given flow and no change in operating cost or in total investment will have occurred and no other component of the plant feels the change.[3] On the other hand, if this transfer of investment results in a net decrease in the stagnation enthalpy leaving the turbine for the given flow rate, then the output per unit of flow would be increased while the total investment remained constant. Obviously, transfer of investment in this sense should continue until no further gain results.

This criterion for balance between the condenser and turbine exhaust is given formal expression as follows

$$dh_2^\circ = \left(\frac{\partial h_2^\circ}{\partial \frac{I_{ex}}{G}}\right)_{I_c} d\left(\frac{I_{ex}}{G}\right) + \left(\frac{\partial h_2^\circ}{\partial \frac{I_c}{G}}\right)_{I_{ex}} d\left(\frac{I_c}{G}\right) = 0$$

where

$$d\left(\frac{I_{ex}}{G}\right) = -\,d\left(\frac{I_c}{G}\right)$$

Let

$$d\left(\frac{I_{ex}}{G}\right) = d\left(\frac{A_{ex}}{G}\right) C_{ex}$$

and

$$d\left(\frac{I_c}{G}\right) = d\left(\frac{A_c U}{G}\right) C_c$$

Then, combining

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)_{\frac{A_c U}{G}} = \left(\frac{\partial h_2^\circ}{\partial \frac{A_c U}{G}}\right)_{\frac{A_{ex}}{G}} \frac{C_{ex}}{C_c} \quad \ldots\ldots\ldots [1]$$

where

A_{ex} = exhaust annulus area, sq ft
A_c = condenser surface, sq ft
$U \equiv \dfrac{G(h_2^\circ - h_0)}{A_c}$; over-all heat-transfer coefficient, Btu/hr sq ft deg F
C_{ex} = investment increment associated with unit change in exhaust annulus,[4] dollars per sq ft
C_c = investment increment associated with unit change in condenser,[4] dollars per Btu/hr deg F
G = steam flow rate to condenser, lb per hr
h_2° = stagnation enthalpy leaving last turbine stage, Btu per lb

---

[1] Associate Professor in Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass. Mem. ASME.

[2] Second Lieutenant, United States Air Force, formerly Junior Engineer, Convair, Fort Worth, Texas.

Contributed by the Power Division and presented at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–21, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 2, 1955. Paper No. 55—S-31.

[3] Strictly speaking, the condensate temperature is affected by a variation in condensing pressure as final stagnation enthalpy remains constant. The effect of this change on capital investment is surely negligible and on operating cost sufficiently small to be neglected.

[4] These marginal rates of investment C_{ex} and C_c include costs of concomitant changes in foundation, structure, circulating-water system, etc.

I_{ex} = investment in exhaust annulus area, dollars
I_c = investment in condensing area, dollars

Equation [1] stipulates that balance between these two components exists when a small increment of investment in either one is equally effective in improving the stagnation enthalpy. It implies that both areas are susceptible to continuous variation. For the condenser this is literally the case; for the turbine it is amazing how closely it is approximated by the large number of exhaust-end arrangements commercially available for central station units of large capacity. Let us, then, consider the effects of small increments of these areas upon the stagnation enthalpy $h_2°$.

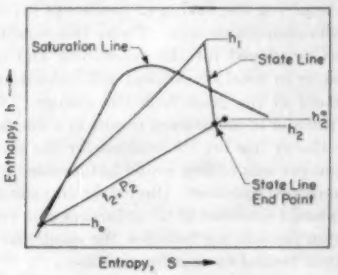


FIG. 1 ENTHALPY-ENTROPY DIAGRAM FOR TURBINE EXHAUST AND CONDENSING PROCESSES

Fig. 1 indicates the exhaust process on the enthalpy-entropy plane. The state line, h_1, s_1 to h_2, s_2, is assumed to be independent, except for its lower terminus, of the condensing pressure. The end-point properties h_2, s_2, v_2, p_2 are all functions of the condensing temperature only. This in turn depends principally on the condenser area because the exhaust loss, $h_2° - h_2$, is small compared to the heat rejected in the condenser

$$(h_2° - h_c) = \frac{A_c U}{G}(t_2 - t_c) \quad \ldots\ldots [2]$$

where

h_c = enthalpy of saturated liquid at temperature t_2, Btu/lb
t_2 = condensing temperature, deg F
t_c = circulating-water temperature, deg F

On the other hand, because of the sensitivity of specific volume to this temperature, the leaving loss is profoundly influenced by the condensing temperature as indicated by the following equation

$$(h_2° - h_2) = (K/2gJ)\left(\frac{v_2}{A_{ex}/G}\right)^2 \times \frac{1}{(3600)^2} \quad \ldots\ldots [3]$$

in which

v_2 = specific volume leaving last stage, cu ft per lb
K = loss coefficient for exhaust hood
g = acceleration of gravity, ft per sec²
J = work equivalent of heat, ft-lb/Btu

Assigning a value of unity to K since it appears that a typical exhaust-hood loss is one axial-velocity head

$$h_2° = h_2 + \frac{\left(\frac{v_2}{A_{ex}/G}\right)^2}{2gJ} \times \frac{1}{(3600)^2} \quad \ldots\ldots [4]$$

Along any one condition curve h_2 and v_2 are uniquely determined by t_2. So

$$h_2° = f\left(t_2, \frac{A_{ex}}{G}\right) \quad \ldots\ldots [4a]$$

However, Equation [2] indicates that since the condensate enthalpy h_c is determined by t_2

$$h_2° = f\left(t_c, t_2, \frac{A_c U}{G}\right) \quad \ldots\ldots [2a]$$

If t_2 is eliminated between Equations [4a] and [2a], there results a relation

$$h_2° = f\left(t_c, \frac{A_c U}{G}, \frac{A_{ex}}{G}\right) \quad \ldots\ldots [5]$$

Since an arbitrary condition curve and the properties of moist steam are involved, Equation [5] has not been reduced to analytic form. However, solutions have been developed by numerical

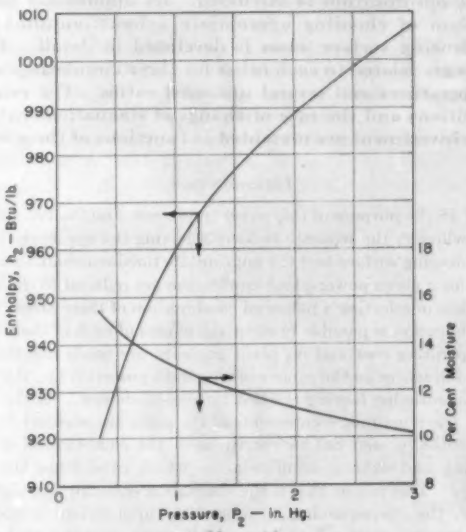


FIG. 2 PROPERTIES ALONG ASSUMED TURBINE STATE LINE

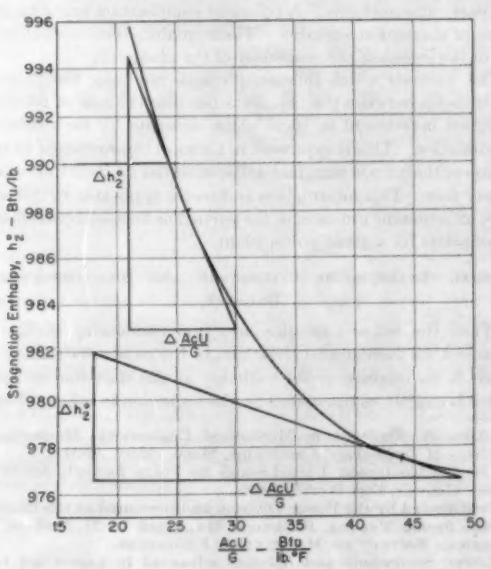


FIG. 3 TYPICAL PLOT OF STAGNATION ENTHALPY AT TURBINE EXHAUST VERSUS CONDENSING AREA
[t_c = 60 F; A_{ex}/G = 1.71 × 10^{-4} sq ft/(lb/hr.)]

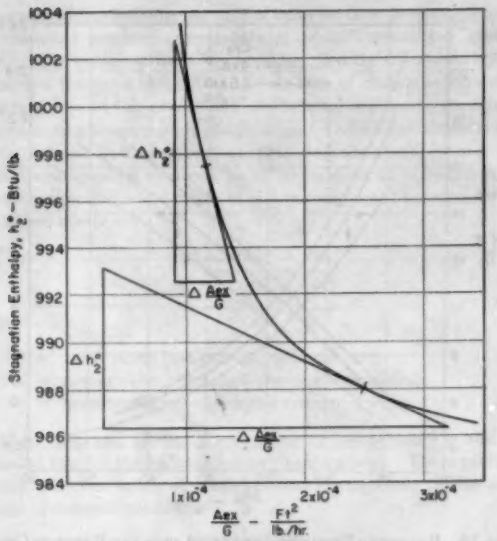


FIG. 4 TYPICAL PLOT OF STAGNATION ENTHALPY AT TURBINE EXHAUST VERSUS EXHAUST-ANNULUS AREA
(t_c = 60 F; A_eU/G = 24 Btu/lb deg F.)

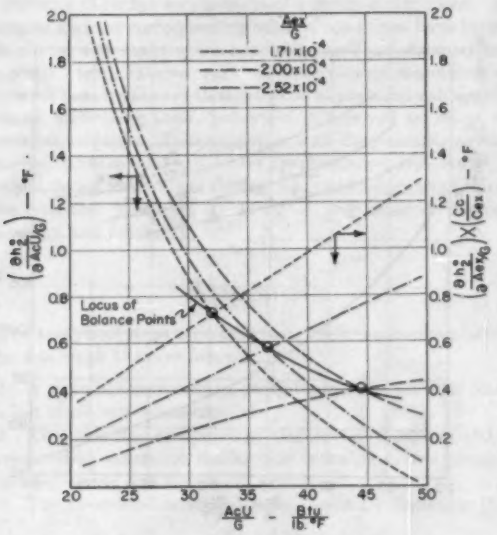


FIG. 5 TYPICAL SUPERPOSITION OF STAGNATION ENTHALPY DERIVATIVES TO DETERMINE BALANCED AREAS
(t_c = 60 F; C_e/C_{ex} = 6.77 × 10⁻⁴.)

methods for t_c equal to 50, 60, and 70 F and a series of A_{ex}/G and A_eU/G values in the range of usual interest. The condition curve chosen is defined by properties plotted in Fig. 2. These correspond to a curve passing through the 12 per cent moisture 1.5-in. Hg point and defined by 2:1 pressure-ratio stages of 80 per cent efficiency.

RESULTS

Figs. 3 and 4 are typical plots of the results of these computations (see Appendix for calculation procedure). In the first instance A_eU/G is taken as the abscissa and A_{ex}/G as the parameter. The second plot inverts the roles of these two variables. These

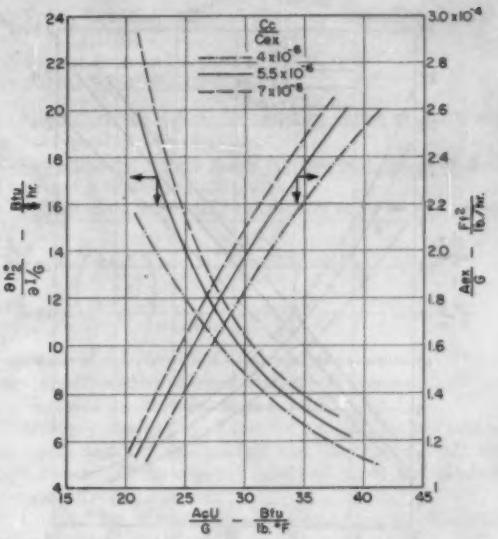


FIG. 6 BALANCED EXHAUST-ANNULUS AREAS AND RATES OF CHANGE OF STAGNATION ENTHALPY WITH PLANT INVESTMENT VERSUS CONDENSING AREA
[t_c = 50 F; C_c = 0.0071 $/(Btu/hr deg F.)]

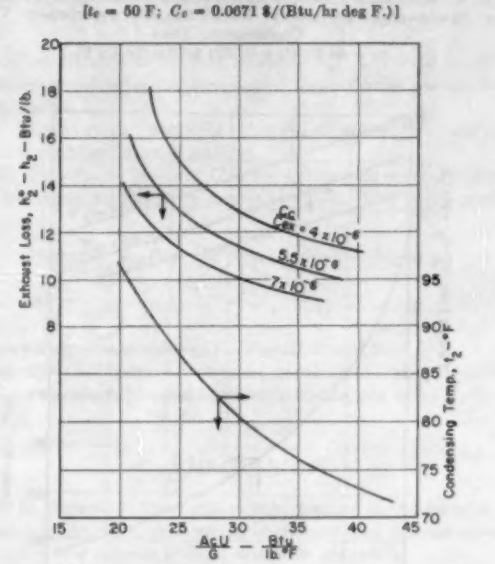


FIG. 7 EXHAUST LOSS AND CONDENSING TEMPERATURE FOR BALANCED AREAS VERSUS CONDENSING AREA (t_c = 50 F)

same figures indicate the graphical differentiation of these curves.

The partial

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_eU}{G}}\right)_{A_{ex}/G}$$

was plotted directly as a function of A_eU/G for several values of A_{ex}/G. The derivatives

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)_{\frac{A_eU}{G}}$$

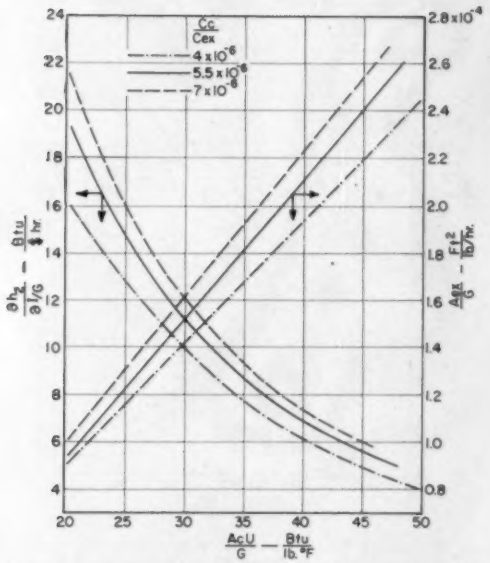


FIG. 8 BALANCED EXHAUST-ANNULUS AREAS AND RATES OF CHANGE OF STAGNATION ENTHALPY WITH PLANT INVESTMENT VERSUS CONDENSING AREA
[t_c = 60 F; C_c = 0.0671 $/(Btu/hr deg F).]

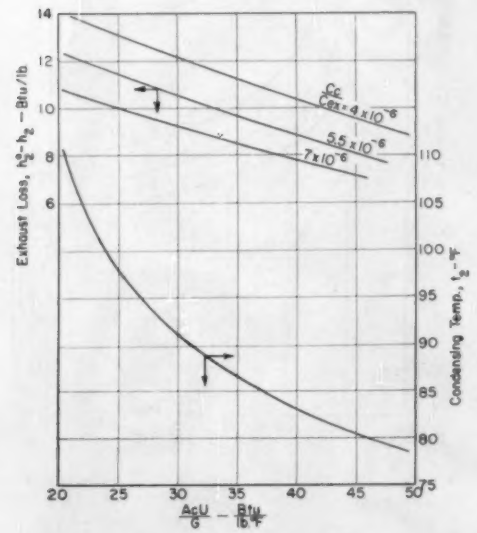


FIG. 9 EXHAUST LOSS AND CONDENSING TEMPERATURE FOR BALANCED AREAS VERSUS CONDENSING AREA (t_c = 60 F)

were multiplied by selected values of the cost ratio, C_c/C_{ex}, and plots of the products were superposed on the

$$\left(\frac{\partial h_2^\circ}{\partial \dfrac{A_c U}{G}}\right)_{A_{ex}/G}$$

plots. A sample of these combined graphs is shown in Fig. 5. The intersections of

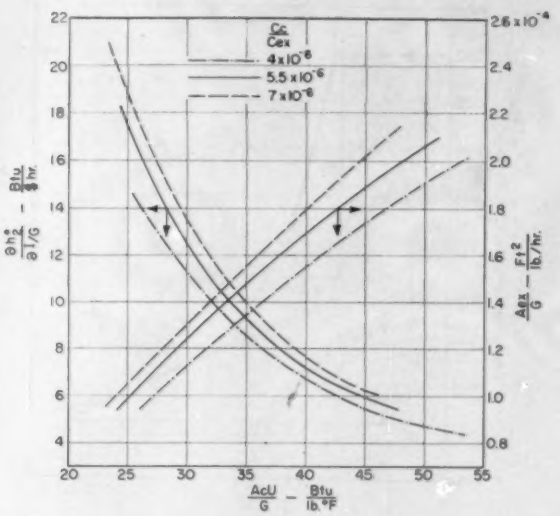


FIG. 10 BALANCED EXHAUST-ANNULUS AREAS AND RATES OF CHANGE OF STAGNATION ENTHALPY WITH PLANT INVESTMENT VERSUS CONDENSING AREA
[t_c = 70 F; C_c = 0.0671 $/(Btu/hr deg F).]

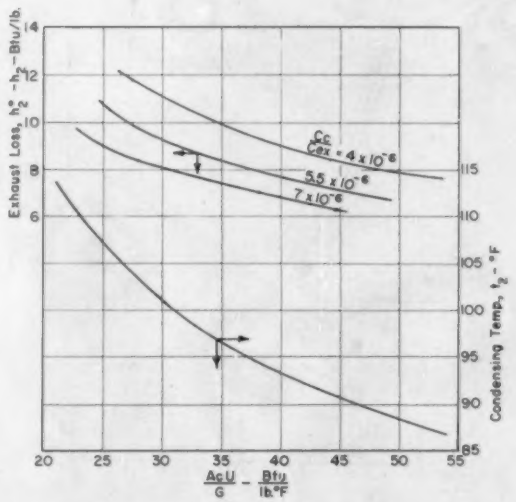


FIG. 11 EXHAUST LOSS AND CONDENSING TEMPERATURE FOR BALANCED AREAS VERSUS CONDENSING AREA (t_c = 70 F)

$$\left(\frac{\partial h_2^\circ}{\partial \dfrac{A_c U}{G}}\right)_{A_{ex}/G} \quad \text{and} \quad \left(\frac{\partial h_2^\circ}{\partial \dfrac{A_{ex}}{G}}\right)_{\frac{A_c U}{G}} \frac{C_c}{C_{ex}}$$

curves for corresponding values of A_{ex}/G constitute solutions of Equation [1]; i.e., they establish balanced values of condenser and exhaust-annulus areas corresponding to the selected values of effective circulating-water temperature and cost ratio.

Figs. 6 and 8 and 10 present A_{ex}/G as a function of A_cU/G and C_c/C_{ex} for effective circulating-water temperatures of 50, 60, and 70 F, respectively. These same figures also present the corresponding values of the stagnation enthalpy $h_2°$ and of its rate of

change with investment per unit of exhaust flow predicated on an incremental condensing-system cost of 6.71 cents per unit of A_eU. The enthalpy itself is, of course, directly influenced by the choice of the condition curve, but its rate of change is somewhat less sensitive to this choice and can be used with some confidence in evaluating the economic justification for additional investment.[5] It is perhaps well to note that, provided one starts from a balanced combination of areas, small increments of investment are equally effective when made in either component because for the balanced condition

$$\frac{dh_2^\circ}{dI} = \frac{\partial h_2^\circ}{\partial I_{ex}} = \frac{\partial h_2^\circ}{\partial I_c}$$

where

I = total plant investment, dollars
I_{ex} = investment in exhaust-annulus area, dollars
I_c = investment in condensing system, dollars

Figs. 7, 9, and 11 plot the condensing temperature t_2 and the exhaust loss for the balanced-area combinations. These plots indicate that the range of areas covered corresponds to the usual range of exhaust conditions.

Conclusions

It has been shown that the problem of balancing condenser and exhaust-annulus areas is substantially independent of the problem of selecting the other components of a steam-power plant. Area balances and the corresponding exhaust conditions have been developed for a typical turbine-condition curve and several circulating-water temperatures and condenser-exhaust-annulus cost ratios. These results are thought to be of practical value and the method underlying their derivation is believed to be of even greater significance. Taken together with the presently available data on turbine and condenser performance and costs[6] this method, or an elaboration thereof, can yield solutions of the condenser-turbine balancing problem of any desired degree of generality and reliability.

Appendix

The analytical steps necessary to develop the results plotted in Figs. 6 through 11 are as follows:

1 A state line representing possible stream conditions leaving the last blade row is constructed.

2 The state-line end-point properties p_2, v_2, and t_2 and the corresponding saturated condensate enthalpy h_0 are plotted or tabulated versus enthalpy h_2.

3 The functional relationship expressed by Equation [5]

$$h_2^\circ = f\left(t_c, \frac{A_eU}{G}, \frac{A_{ex}}{G}\right)$$

is determined point by point by the following:

(a) Assigning values of A_{ex}/G and t_c.
(b) For each of a series of values of h_2 calculate h_2° from Equation [4].
(c) From Equation [2] calculate corresponding values of A_eU/G.
(d) Repeat for several values of A_{ex}/G.
(e) Repeat for additional values of t_c.

4 Plot h_2° versus A_eU/G for constant values of A_{ex}/G and t_c (Fig. 3 is typical of such plots).

5 Cross-plot h_2° versus A_{ex}/G for constant values of A_eU/G and t_c (Fig. 4 is typical of such plots).

6 Differentiate graphically the plots of steps 4 and 5 to establish

$$\left(\frac{\partial h_2^\circ}{\partial \dfrac{A_eU}{G}}\right)_{A_{ex}/G,\,t_c} \quad \text{and} \quad \left(\frac{\partial h_2^\circ}{\partial \dfrac{A_{ex}}{G}}\right)_{\frac{A_eU}{G},\,t_c}$$

as functions of A_eU/G with A_{ex}/G and t_c as parameters. This procedure is facilitated by differentiating the h_2° versus A_{ex}/G curves at the values of A_{ex}/G chosen in step 3 (a).

7 Multiply the values of $(\partial h_2^\circ)/(\partial A_{ex}/G)$ by the unit cost ratio C_e/C_{ex} and plot this product and $(\partial h_2^\circ)/(\partial A_eU/G)$ versus A_eU/G for each of the selected values of A_{ex}/G and constant t_c (for example, Fig. 5).

8 Connect the intersections of corresponding A_{ex}/G curves of step 7 to yield

$$\frac{\partial h_2^\circ}{\partial \dfrac{A_eU}{G}} = \left(\frac{\partial h_2^\circ}{\partial \dfrac{A_{ex}}{G}}\right)\frac{C_e}{C_{ex}} \text{ versus } \frac{A_eU}{G} \text{ (see Fig. 5)}$$

These are balance points and establish A_eU/G versus A_{ex}/G for balance corresponding to the specified cost ratio and cooling-water temperature.

9 Repeat steps 7 and 8 for other values of cost ratio and cooling-water temperature of interest.

10 Plot A_{ex}/G versus A_eU/G for balance for several values of cooling-water temperature and cost ratio. Plot corresponding values of

$$\left(\frac{\partial h_2^\circ}{\partial \dfrac{A_eU}{G}}\right)C_e = \frac{\partial h_2^\circ}{\partial \dfrac{I}{G}}$$

for any appropriate value of C_e (Figs. 6, 8, and 10).

11 For each balanced relationship calculate such quantities of interest as exhaust loss, condensing pressure, etc. (Figs. 7, 9, and 11).

Discussion

N. R. Deming.[7] The authors have treated the problem of determining the most economical condenser and turbine-exhaust sizes in a new manner with an academic approach.

However, it appears that there are practical difficulties in applying the method and aspects of the over-all problem which were not considered. Such items as load factors, fuel costs, and seasonal water-temperature variations affect the selection of areas as well as condensing-water temperature and turbine and condenser costs. Even though a large number of exhaust sizes may appear to be available for one turbine rating, in most cases only two would be considered and these may vary by 25 per cent from each other. Since the exhaust-area increments are large and the cost ratio is over 1000 to 1, condenser surface has little effect on the selection of exhaust size.

It would appear that a similar method of analysis could be applied to determine for a given exhaust size the condenser surface

[5] This statement could be checked by repeating this analysis for several condition curves with the result that final moisture, for example, would become another parameter. The authors ask for the reader's indulgence with respect to the omission of this step.

[6] For example, "The Economics of Large Reheat Turbine-Exhaust-End Size Selection," by D. W. R. Morgan, Jr., and S. D. Fulton, Trans. ASME, vol. 77, 1955, pp. 363–371.
"Economic Determination of Condenser and Turbine-Exhaust Sizes," by E. H. Miller and A. Sidun, Trans. ASME, vol. 77, 1955, pp. 373–383.

[7] Design Engineer, Steam Engineering Department, Westinghouse Electric Corporation, Philadelphia, Pa. Assoc. Mem. ASME.

which would produce the lowest stagnation enthalpy for a given set of assumptions.

E. H. MILLER.[8] The authors have presented an interesting analytical method for matching condensers and turbine-exhaust sizes.

The concept of transferring increments of investment between turbine annulus and the condensing system in order to obtain the most output for a given investment appears to be valid and useful to those interested in sizing and selecting station components.

The paper would have been more useful if an example of its practical application had been included plus more details on the selection of cost ratio C_c/C_{ex} and the heat-transfer coefficient U, used in deriving the practical curves in Figs. 3 to 11. For instance, the writer obtained a value of approximately 8.2×10^{-4} for C_c/C_{ex} using the authors' references, compared to values of C_c/C_{ex} of 5.5×10^{-4} shown in Figs. 6 to 11.

The authors' analytical approach should yield a very precise balance of areas for the assumed conditions; however, it appears difficult to factor in the many operating conditions with which the station designer is faced, and to account for changes in turbine performance for design changes that accompany changes in turbine-annulus areas.

For these reasons it would appear desirable to continue using availabile numerical methods which are easy to apply, and take into account the actual station-performance changes of turbines and condensers as they are built and operated.

A. K. SMITH.[9] The concept of economic balance as related to power-plant optimization introduced by the authors is an interesting and significant development which should prove to be of great assistance to the power-plant designer. This concept of balanced conditions existing in power plants of optimum economic design and the methods of expressing the criterion in concise analytical form as demonstrated by this paper are welcome contributions to the art of power-plant design. In addition to presenting balanced areas for three circulating-water temperatures and several unit-cost ratios, the authors present the rate of change of stagnation enthalpy with investment as a function of these areas in a form that is most convenient for economic optimization of power plants.

In the introduction and formal analysis of condenser area and exhaust-annulus-area balance the authors develop as the criterion of balance the equation

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)_{\frac{A_c}{G}} = \frac{C_{ex}}{C_c}\left(\frac{\partial h_2^\circ}{\partial \frac{A_c}{G}}\right)_{\frac{A_{ex}}{G}} \quad \ldots\ldots\ldots\ldots [1]$$

Expressions necessary for the evaluation of the functions related by this equation are clearly developed by the authors on reasonable and acceptable premises. The authors then proceed to develop area balances for typical conditions in order to present results of practical value. It can be verified that the series of analytical steps 1 to 7 given by the authors in the Appendix of the paper will yield the data necessary to present

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)\frac{C_c}{C_{ex}} \quad \text{and} \quad \left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right)$$

as functions of $\left(\frac{A_cU}{G}\right)$ with $\left(\frac{A_{ex}}{G}\right)$

as a parameter for a constant value of (t_c) as shown in Fig. 5 of the paper. It is the contention of the authors that the intersections of the curves for corresponding values of (A_{ex}/G) and a given (t_c), as shown in Fig. 5, constitute solutions of Equation [1] and that they establish balanced values of condenser areas and exhaust-annulus areas corresponding to the selected value of circulating-water temperature (t_c) and cost ratio (C_c/C_{ex}).

Since Equation [1], the criterion for economic balance, contains the term

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_c}{G}}\right)$$

and Fig. 5 involves the term

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right)$$

the writer believes the expression of Equation [1] in terms of

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right)$$

to be in order.

Divide both sides of Equation [1] by U to obtain

$$\frac{1}{U}\left(\frac{\partial h_2^\circ}{\partial \frac{A_c}{G}}\right) = \left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)\frac{C_c}{UC_{ex}} \quad \ldots\ldots\ldots\ldots [6]$$

It can be stated that

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_c}{G}}\right) = \left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right)\left(\frac{\partial \frac{A_cU}{G}}{\partial \frac{A_c}{G}}\right) \quad \ldots\ldots\ldots\ldots [7]$$

and that

$$\left(\frac{\partial \frac{A_cU}{G}}{\partial \frac{A_c}{G}}\right) = U \quad \ldots\ldots\ldots\ldots [8]$$

It follows that

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_c}{G}}\right) = \left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right)U \quad \ldots\ldots\ldots\ldots [9]$$

and that

$$\frac{1}{U}\left(\frac{\partial h_2^\circ}{\partial \frac{A_c}{G}}\right) = \left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right) \quad \ldots\ldots\ldots\ldots [10]$$

Therefore

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_cU}{G}}\right) = \left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)\frac{C_c}{UC_{ex}} \quad \ldots\ldots\ldots\ldots [11]$$

---

[8] Turbine Advance Engineering, Large Steam Turbine-Generator Department, General Electric Company, Schenectady, N. Y. Assoc. Mem. ASME.

[9] Assistant Engineer, System Planning Division, Engineering Department, New Orleans Public Service, Inc., New Orleans, La.

Equation [11] is equivalent to Equation [1]. Comparison of this equation with the graphical solution presented in Fig. 5 of the paper indicates that the coefficient U is absent from the term

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)\frac{C_e}{C_{ex}}$$

The dimensions of the term

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_{ex}}{G}}\right)\frac{C_e}{C_{ex}}$$

are Btu/hr sq ft and not deg F as is indicated in Fig. 5. The dimensions of the term

$$\left(\frac{\partial h_2^\circ}{\partial \frac{A_eU}{G}}\right)$$

are deg F as is indicated in Fig. 5. The fact that these terms are not dimensionally homogeneous raises a question as to the validity of the contention of the authors that Fig. 5 constitutes a solution to the problem of obtaining balanced areas. Furthermore, the graphical solution indicated in Fig. 5 is not equivalent to the solution of Equation [1].

The authors state that the results as presented in Figs. 6, 8, and 10 are predicted on an incremental condensing-system cost of 6.71 cents per unit of (A_eU). This suggests, perhaps, that the term (C_e) in Fig. 5 is not defined to be the condensing-system cost per unit (A_e) of Equation [1], but rather the condensing-system cost per unit (A_eU). This change in definition will remove the question of the validity of the authors' contention and make the solution of Fig. 5 a solution of Equation [1].

If the authors will explain steps 7 and 8 of the procedure and Fig. 5 in greater detail this question can be resolved and the noteworthy results of this paper can be established firmly.

Authors' Closure

An apology is due those readers who, like Messrs. Smith and Miller, have run afoul of a confusion in the definition of condenser cost as included in the preprints of this paper. As Mr. Smith has surmised the proper definition is

$$C_e = dI_e/d(A_eU)$$

This correction together with another relating to the definition of U (see p. 135) should clarify the dimensional difficulties cited by Mr. Smith as well as the question of choice of representative C_e/C_{ex} values as raised by Mr. Miller.

Mr. Deming notes some very real limitations on the use of this analysis. It must be confessed that if it is not possible to assign an economically representative circulating-water temperature the method has little to offer. In this case there is no recourse except a brute-force analysis of many combinations of components subjected to the entire gamut of operating conditions. On the other hand, the items of load factor and fuel cost which are mentioned relate, as does the rate of fixed charges, to the selection of the particular balanced combination which is appropriate. These factors do not affect the balancing operation itself.

With respect to Mr. Deming's comment on the relative costs of condenser and annulus areas it is felt that the results of the analysis speak for themselves. There appears to be no justification for calling a cost ratio of 1000 to 1 large, because it is clearly dependent on the units in which these areas are measured. Since they have entirely different functions, direct comparision of their per-square-foot costs has little meaning.

It must be acknowledged that exhaust annuli can only be varied as a practical matter in finite increments. However, recent price lists indicate the availability of some eight different exhaust-end areas for 150 MW units. From this listing a remarkably linear relationship between area and turbine cost can be deduced. The method, it would seem, has particular application to narrowing the range within which the final selection must be made. In most instances, of course, *the* balanced combination selected on the basis of this analysis must be compromised to the available areas. After this has been accomplished the well-established numerical techniques of performance prediction as suggested by Mr. Miller must be applied. It is the authors' feeling, however, that these latter methods alone do not reveal the influence of unit costs and water temperatures on the interactions between the area choices as lucidly as do the results of the analysis presented.

The need for an example to illustrate the application of the method is felt by the writer, too. It was deliberately omitted in the present instance because of the large number of considerations involved which are extraneous to the central idea of the paper. A paper dealing more broadly with the optimization problem and including an example is now in preparation.

Improving Airplane Handling Characteristics With Automatic Controls

By C. L. MUZZEY,[1] BUFFALO, N. Y.

The paper describes an important mode of airplane flight in which the aircraft is under the joint control of a human pilot and an automatic control system. The first part of the paper describes the interaction of air loads and inertia reactions which govern the dynamic stability of the airplane, and seeks to give a physical picture of the response of the airplane to control and random disturbances. The paper then describes a research project in which the pilot's controls were disconnected from the control surfaces, and separate servos provided both to drive the aerodynamic surfaces, and to resist the actions applied to the cockpit controls by the pilot. The surface servos receive signals from cockpit-control motions and from a stability computer which senses airplane-body motions. Airplane response as seen by the pilot can be varied by adjusting the computer, and the pilot is unaware of the control-surface motions caused by the computer. Thus both airplane stability and control feel can be varied independently and easily in flight. Research airplanes so equipped are now being used to map out regions of acceptability for future design in the handling characteristics problem, and are incidentally demonstrating means for stabilizing existing airplanes which may be deficient in one or more respects.

Introduction

THE modern high-speed airplane is one of the most important users of automatic control systems. They are found all the way from small instrument repeater servos, through engine power, temperature, or thrust controls to fully automatic flight-control systems which direct the path of the airplane or missile with no human help whatsoever during part or all of its mission.

This paper deals with a very important mode of aircraft operation in which the airplane is under the joint control of a human pilot and an automatic servocontrol system. The ultimate purpose of such a configuration is to relieve the human pilot of the task of supplying routine airplane stabilizing action through his controls by entrusting these duties to the automatic control system. In many airplanes which fly at high speeds, the natural frequencies of motion are so high, on the order of 1 cycle per sec (cps), that the human pilot no longer can act quickly enough to damp them. Fig. 1 shows a block diagram of an airplane so equipped. In so far as the pilot is concerned, the net response of the array of units within the dotted line is the airplane response, and the action supplied by the automatic control system has come to be known in the aircraft industry as "artificial stability" in contrast to the natural stability of the air frame alone.

In order to close the loop of the automatic control system successfully and in order to understand better what stability deficiencies should be improved, we will first try to develop a physical understanding of the black box marked "air frame."

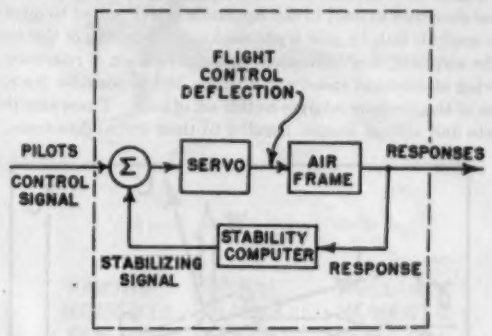


Fig. 1 Airplane Equipped With Artificial Stability

Stability and Control of Typical Airplanes

Historical Development. An airplane in flight is assumed for stability-analysis purposes to be a rigid body, and because its motion is in no way determined by guide rails, axes of rotation, or the like, its behavior is entirely the result of the aerodynamic and gravitational forces and moments acting upon the inertia characteristics of the body. Euler wrote the general equation of motion governing this problem some time in the eighteenth century, but these equations were far too complicated to apply directly to the airplane when it appeared on the scene. Other investigators took advantage of the special conditions attending the flight of an airplane, its symmetry, and its small departures from equilibrium during many practical maneuvers and, early in the twentieth century, Bryan and Bairstow, and Williams in Great Britain presented a set of linearized equations of motion for the airplane in which motions in the plane of symmetry could be separated from those out of the plane of symmetry. These are known as the longitudinal and lateral motions, respectively. We will first look at the longitudinal motions.

Longitudinal Motion. The longitudinal equations of motion are presented in the Appendix, Equations [13], [14], and [15]. They are dimensional differential equations expressing, respectively, the balance of drag forces, lift forces, and pitching moments. Fig. 17 in the Appendix shows the system of moving axes which are used in developing the equations of motion. The axes originate at the center of gravity of the airplane and their direction is fixed with respect to the airplane structure. For quantitative stability analysis, two major advantages result from the use of axes fixed in the airplane: (*a*) The inertia characteristics are constants and (*b*) recording and control instruments measure airplane motions with respect to these same axes.

If transfer functions are required, the three equations of motion may be solved simultaneously to give the relations of fractional forward speed u/U_0, angle of attack α, and pitch angle θ, to the forcing function, elevator angle δ_e.

The determinant of the coefficients of the left-hand side of Equations [13], [14], and [15] yields a fourth-order expression in

[1] Cornell Aeronautical Laboratory.

Presented at the Instruments and Regulators Division Conference, Ann Arbor, Mich., April 25–26, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 4, 1955. Paper No. 55—IRD-3.

terms of the operator D. This expression is the denominator of the longitudinal transfer functions, and when equated to zero is called the characteristic equation of the dynamic system. A numerical solution for roots of the quartic typically show two damped oscillatory modes with well-separated natural frequencies. For the C-45 Beechcraft airplane flying at 150 mph at 5000 ft, the long period (phugoid oscillation) has a period of 35 sec while the more rapid motion has a period of 2 sec. The wide separation in frequency of the two modes makes it possible to discuss and analyze them separately.

Phugoid Oscillation. The moving-axis system fixed to the airplane structure as used in the Appendix is well suited to quantitative analysis but, to gain a physical understanding of the motion of the airplane, it is convenient to think of a set of reference axes moving at constant speed and altitude and to consider the excursions of the airplane relative to this set of axes. These axes do not rotate and always remain parallel to their initial directions.

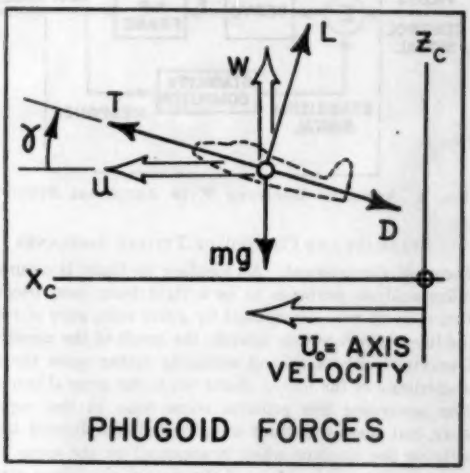


FIG. 2 PHUGOID FORCES

Fig. 2 shows the airplane in relation to the constant-velocity axes which move from right to left at the steady velocity U_o. The axes are designated X_e and Z_e to distinguish them as constant-velocity axes. Superimposed on the steady axis velocity are the incremental velocities of the airplane u and w in the X_e and Z_e-directions. The airplane will have thrust T, and drag D, along the resultant direction of motion, and lift L at right angles to the direction of motion. The path inclination γ is also defined. Lift and drag are permitted to vary as the square of the changing flight velocity, while the lift and drag vectors tilt slightly as the path angle changes. The lift vector casts a component along the X-direction as it tilts back and forth. The resulting motions of the airplane center of gravity are found to be slow and lightly damped oscillations vertically and horizontally about the reference point. The natural frequency is found to depend inversely on the equilibrium speed U_o, while the damping ratio of the motion depends on the difference between drag and thrust characteristics divided by the lift coefficient. Note that the rotational dynamics of the airplane have not been considered in this analysis. Actually, they modify slightly the natural frequency and damping.

The same results are obtained either by solving for the X_e and Z_e motions of the airplane CG with respect to the constant-velocity axes of Fig. 2, or by solving simultaneously Equations [1] and [2] of the Appendix, and neglecting the α-terms. In either case, the motion of the airplane CG appears as an elliptical path around the moving reference point. Damping brings the path spiraling into the origin as shown in Fig. 3.

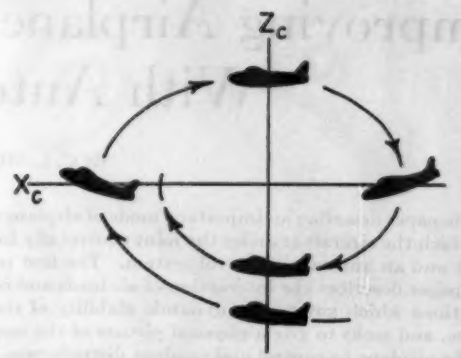


FIG. 3 PHUGOID MOTION

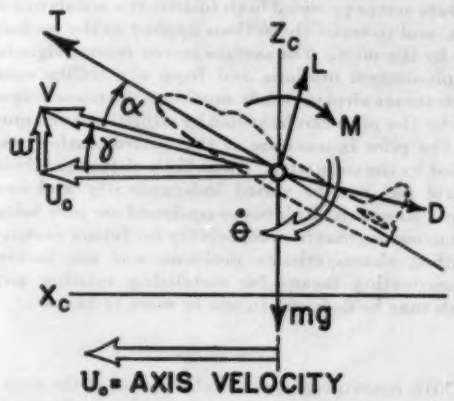


FIG. 4 SHORT-PERIOD MOTION

The solution of Equations [1] and [2] is carried out in the Appendix in the form of a transfer function relating change in forward speed to a lift increment as the forcing function.

The pilot must contend with this oscillation whenever he wants to set his speed or altitude, and it can be seen that he excites the phugoid directly when he changes his thrust control, or when he changes his elevator setting, which in turn changes the equilibrium relation between lift L and velocity U. A vertical or a horizontal gust also will disturb the balance of forces and excite the phugoid oscillation. A flight-test record of this oscillation will be shown later.

Short-Period Oscillation. The short period can be visualized as the vertical and rotational motions of the airplane, again, with respect to a set of axes traveling in a straight line with uniform velocity. In this case, however, there is no incremental forward velocity with respect to the axes, but the pitching rotations are rapid, and the moment of inertia in pitch and the aerodynamic pitching moments must be included. Fig. 4 shows the geometry and forces and moments acting in the short-period oscillation. The vertical Z_e-force and the pitching-moment equations can be written and solved indicating a well-damped oscillation. Again, for illustrating the motion and forces and moments, the constant-velocity nonrotating axes are shown as in the previous example. Note that the thrust-force and drag-force directions now differ by the angle α which was zero for the long-period oscillation. Solution of the simplified problem leads to the same conclusions

as the more complete solution in the Appendix. The natural frequency depends on the indicated air speed directly, the aerodynamic rotational stiffness, and inversely on the moment of inertia and mass. The damping arises from both rotational damping and vertical motion damping.

When the pilot changes the elevator setting he applies a pitching moment to the airplane, and the angle of attack adjusts to a new value according to the dynamics of the short period. A gust, especially a vertical gust, will excite the short period by changing the angle of attack. The character of the short-period response is shown in Fig. 5 in the early part of the vertical acceleration-time history of the C-45 following an elevator step deflection.

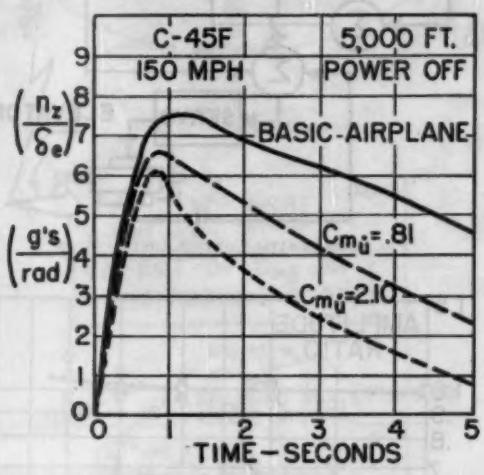


Fig. 5 Normal Acceleration Response

Lateral Equations of Motion. The lateral motions of an airplane also can be linearized and simplified, and consider the small departures of the airplane from an equilibrium path. The motions taken into account are rolling, sideward displacement, and heading change or yawing. The lateral equations of motion are given in the Appendix, Equations [30], [31], and [32]. When the three equations are solved, they result in a fourth-order system which in most cases exhibits two real roots, one strongly negative, the other either slightly positive or slightly negative, and a complex pair.

The large negative root merely says that when a pure rolling moment is applied, the airplane quickly reaches a steady rolling velocity in an exponential manner. This root incidentally becomes positive with the airplane stalled.

The complex pair of roots represents an oscillation with a period of 2.4 sec for the C-45 and a damping ratio of 11 per cent. The motion is a combination of rolling and yawing called the Dutch roll. The airplane has a weather-vane stability about its vertical axis and if restrained to this axis will perform damped oscillations when disturbed. When it is free to roll, as well, the dihedral effect of the wings causes an aerodynamic rolling moment every time the plane sideslips. As a result the airplane rolls and, in turn, the rolling velocity gives rise to aerodynamic yawing moments. Phase relations are such that the rolling motion decreases the damping of the Dutch roll. Fig. 6 illustrates the Dutch-roll motion approximately, assuming that the center of gravity follows a straight path in space.

The third lateral mode is called the "spiral" convergence or divergence depending on its nature. In this mode the path followed by the airplane center of gravity does curve while the sideslip angle is determined by the balance of moments which

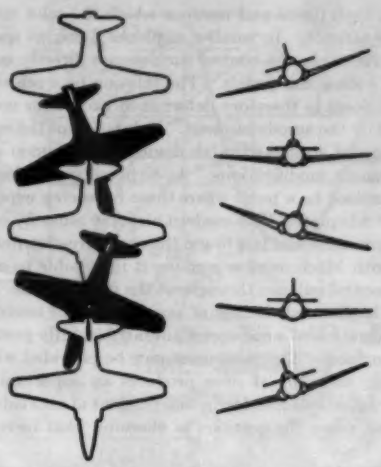


Fig. 6 Dutch-Roll Motion

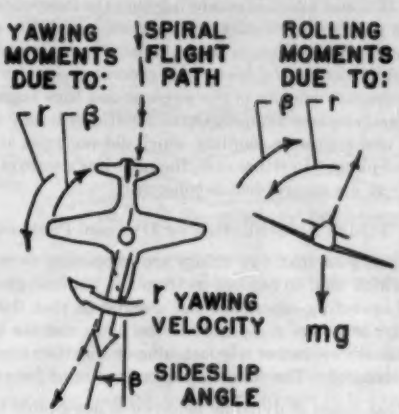


Fig. 7 Spiral Stability

change so slowly that the yawing inertia moments are negligible. Fig. 7 shows an airplane slowly diverging away from a straight path, with the angular velocity, sideslip, and the corresponding moments acting on it. If the airplane sideslips enough, as shown, the dihedral effect will roll the airplane and head it back toward the straight path. If the destabilizing rolling moment due to yawing velocity predominates, the airplane will diverge. The spiral divergence is so slow that there is no problem of the human pilot not being able to stabilize it easily and usually unconsciously. The problem is rather that he must never divert his attention for any length of time lest the airplane turn away from its proper course and pick up speed and lose altitude in a divergent manner.

Unconventional Designs. Modern trends in airplane design are all acting to reduce the damping of its motions and to increase the natural frequencies of the rapid oscillations. The density of airplanes is increasing while they are flying at higher altitudes where air density is low. Also, wings are becoming smaller, and the reduction in wing span reduces the aerodynamic damping in roll, while high Mach numbers of flight move the wing center of pressure aft thus increasing the pitching-spring restoring rate and raising the natural frequency of the short-period oscillation.

Control Feel. Control feel in our discussion is concerned gener-

ally with both forces and motions which the pilot must exert in flying his aircraft. In smaller airplanes flying at speeds on the order of 200 mph, the control surfaces are directly connected to the pilot's stick and pedals. The relation between stick motion and stick force is therefore determined except for certain tricks practiced by the aerodynamicist. Relocation of the surface hinge line and use of special trim-tab devices are common examples of hinge-moment modifications. As airplane speeds increased, the loads increased to a point where these balancing expedients were no longer adequate. The modern airplane must fly over a range of Mach numbers and it is found that pressure-distribution shapes change with Mach number making it impossible to obtain satisfactory control balance throughout the flight régime. This problem finally dictated the use of power-operated controls and the latest generation of sonic-speed aircraft has fully power-operated control surfaces. The pilot must now be provided with artificial stick feel; and this at once produces an opportunity to make his stick deflections and forces independent of air loads but, at the same time, raises the question of choosing what forces he should have.

The physical separation of the control surfaces from the pilot's stick has introduced another important opportunity for the designer. It is now possible to add together the independent signals from the pilot's control stick and from some stability-computing device and feed this sum to the servo driving the control surface. The stability-computing device may generate a signal proportional to some angular velocity of the airplane and thus augment a certain natural airplane damping characteristic, or it may produce an entirely new synthetic coupling which did not exist at all in the normal airplane. In either case, the pilot is unaware of the action in so far as his control feel is concerned.

Flight Investigation of Dynamic Problems

We have seen that two things are happening to modern airplanes which tend to remove us from the familiar ground where years of operating experience had assured us that the dynamics of the airplane were acceptable to the pilot, and the stick forces and motions were not overly fatiguing or did they invite dangerous overcontrol. The following questions must be answered:

1 What range of dynamic behavior is acceptable to the pilot before he must have artificial help?

2 Having added the devices for artificially helping the pilot, what are the most desirable settings for such apparatus?

3 Having relieved the pilot's control of the aerodynamic loads, what artificial feel is most desirable?

Several research projects at Cornell Aeronautical Laboratory have attacked these questions under sponsorship of both the Navy and Air Force. One of the most interesting was recently carried out under the direction of Mr. Edwin Kidd of the Flight Research Department on a C-45 airplane for the All-Weather Flying Section, Flight Test Division, WADC.

C-45 Project. The control system of a C-45 airplane was modified by installing a hydraulic servo on each of the three flight-control surfaces, elevator, rudders, and ailerons. The pilot's controls were then made so they could be disengaged by spline mechanisms from the servo-operated control surfaces, and three more hydraulic servos were connected to the pilot's controls to provide his artificial feel. Fig. 8 shows a schematic diagram of the elevator-system installation. The elevator to the right of the picture is driven by a hydraulic positioning servo which matches the signals from the stability computer and pilot's stick against a position feedback signal from the servo itself. The elevator motion changes the airplane motion, as shown earlier, and these changes are noted by the stability computer which takes the desired action.

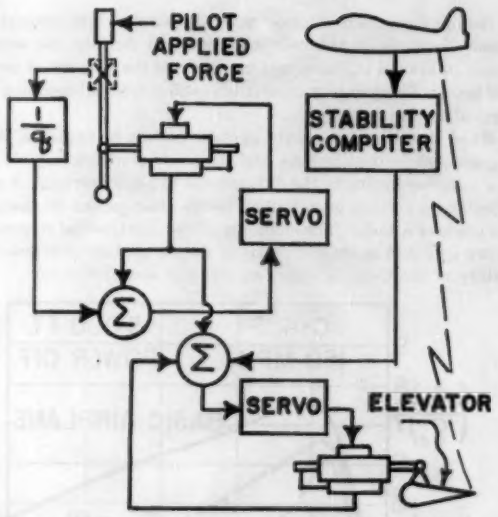


Fig. 8 Schematic Control Installation

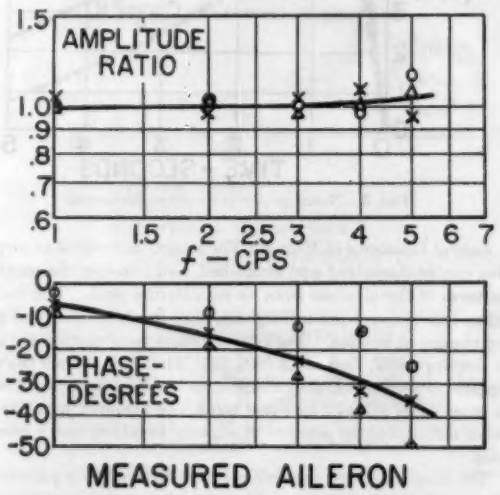


Fig. 9 Measured Aileron Servo Response

The artificial feel system appears on the left-hand side of the diagram and is built around the control stick which is moved by a hydraulic positioning-type servo. The servo moves the stick until its position feedback signal matches the signal derived from the pilot applied force. The harder the pilot pushes, the farther the stick moves, thus simulating the natural aerodynamic restoring forces. In a manual system, the restoring rate of the air loads increases with the dynamic pressure of the air stream, reducing the stick motion for a given force at high speed. This action is duplicated by dividing the electrical force signal by q, the dynamic pressure. The stick position as measured by the feedback potentiometer is fed as a signal to the elevator servo system.

The servos have hydraulic actuators controlled by electrohydraulic servo valves. Feedback and input signals are electric, permitting maximum ease of switching, adding, filtering, amplifying, and other computing such as dividing by dynamic pressure. Fig. 9 shows the frequency response of the aileron servo.

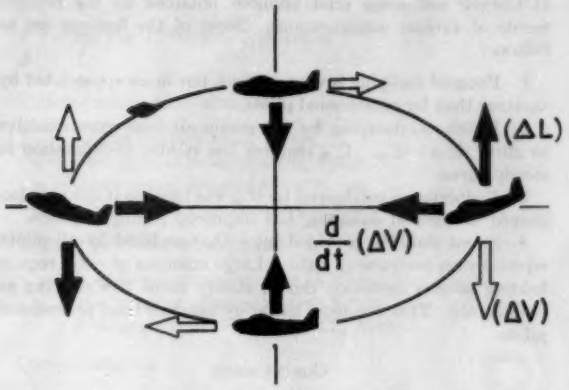


FIG. 10 PHUGOID DAMPER

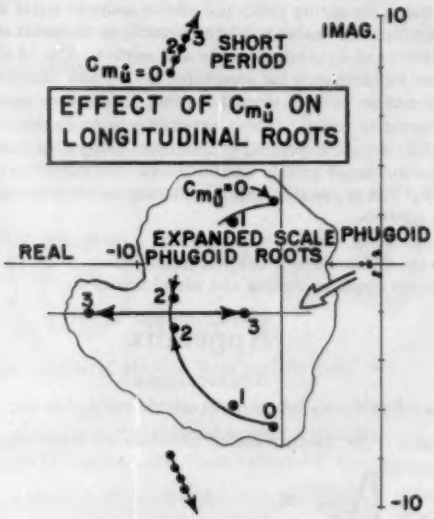


FIG. 11 EFFECT OF $C_{m_{\dot{u}}}$ ON LONGITUDINAL ROOTS

The Stability Computer. Technically, the design of the stability computer is the most interesting part of the installation because the over-all dynamic system includes the airplane as one block in the closed loop. The goal in mind, in providing this artificial stability, was to make all the oscillations of the airplane aperiodic if possible.

Calculations showed that the phugoid oscillation had an undamped natural frequency of 0.028 cps (or undamped period of 35.3 sec) and a damping ratio of 0.081. The damping of the short-period oscillation was found to be 74 per cent for the same flight condition and this was felt to be adequate.

Control of the phugoid was to be made through the elevator, and it will be recalled that moving the elevator causes the airplane to trim at a new angle of attack, and at a given air speed this changes the wing lift. The damping is accomplished by sensing air speed, computing its first derivative, and deflecting the elevator in proportion. The required sensitivity was calculated to be 0.82 deg of up elevator for 1-mph/sec rate of increase of air speed. This amounts to creating a synthetic stability derivative $C_{m_{\dot{u}}}$, pitching moment due to air-speed acceleration, which does not exist at all normally in the airplane.

Fig. 10 shows the path of the airplane during a phugoid oscillation, and shows that the incremental lift caused by elevator motion opposes the downward velocity component of the airplane and thus produces damping. Fig. 11 shows the calculated effect on the roots of the stability quartic of changes in the value of the derivative $C_{m_{\dot{u}}}$. As the derivative increases from 0, the imaginary roots diminish until at $C_{m_{\dot{u}}} = 2.1$ they vanish and two real negative roots appear. There is a small percentage effect on the short-period roots and this also is shown.

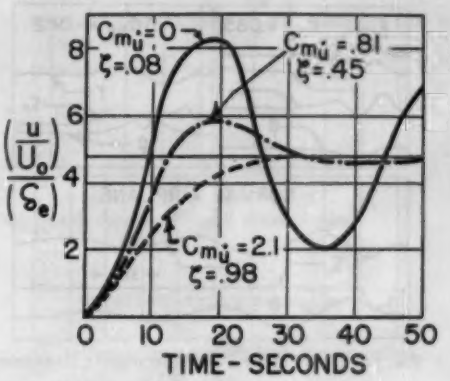


FIG. 12 C-45 ELEVATOR STEP RESPONSE

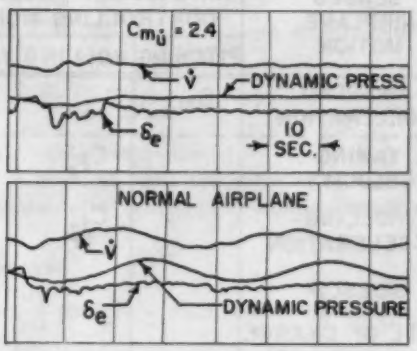


FIG. 13 C-45 FLIGHT-TEST PHUGOID RESPONSE

Fig. 12 shows calculated responses of air speed and vertical acceleration following a step-function elevator deflection for the standard airplane and for two values of $C_{m_{\dot{u}}}$. Finally, Fig. 13 shows flight-test records of dynamic air pressure and elevator deflection for the normal airplane and for $C_{m_{\dot{u}}} = 2.4$.

Lateral Stability. The lateral stability was attacked also by evaluating the effect of changing artificially the values of the existing stability derivatives of the airplane, both aerodynamic and inertial. An analog-computer study also was carried out, and the effect of certain synthetic derivatives evaluated and eventually incorporated into the system. Space does not permit a complete description of the lateral artificial stability system; however, the effect of two of the artificial lateral stability derivatives, one augmenting natural damping, the other synthetic, are shown in flight-test records in Fig. 14. Here the Dutch-roll response to rudder kicks is presented for the normal airplane and with $\Delta C_{n_r} = -0.042$ and $C_{n_{\dot{\beta}}} = 0.085$. The damping ability of either is evident.

Fig. 15 shows the complete list of artificial stability derivatives used in the C-45 project.

The stability computer and the control-surface servo act together ideally to increase the value of a natural stability deriva-

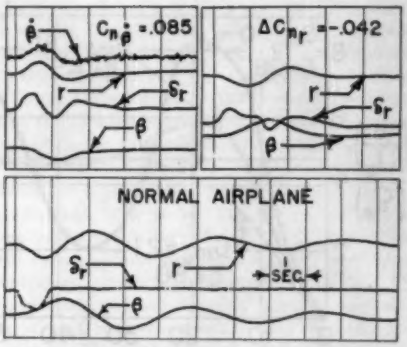


FIG. 14 C-45 FLIGHT-TEST DUTCH-ROLL RESPONSE

SENSED AIRPLANE MOTION	ARTIFICIALLY PRODUCED CONTROLLING MOMENT		
	PITCHING	ROLLING	YAWING
AIRSPEED ACCELERATION	$C_{m_{\dot{u}}}$		
YAWING VELOCITY		C_{l_r}	C_{n_r}
ROLLING ACCELERATION		$C_{l_{\dot{p}}}$	
SIDESLIP			C_{n_β}
RATE OF CHANGE OF SIDESLIP			$C_{n_{\dot{\beta}}}$
YAWING ACCELERATION			$C_{n_{\dot{r}}}$

FIG. 15 ARTIFICIAL STABILITY FUNCTIONS FOR C-45

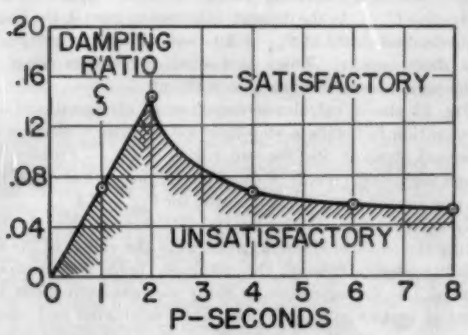


FIG. 16 REQUIREMENTS FOR DUTCH-ROLL DYNAMICS

tive, or to create a synthetic derivative. If the computer or the servo has appreciable phase lag at frequencies near the natural frequencies of the airplane or at frequencies at which the airplane must be maneuvered, the out-of-phase component of the artificial derivative actually will modify some other motion of the airplane. Such effects must be compensated by another appropriate artificial derivative or by using servos and computer elements with very small phase lags up to 4 or 5 cps.

Test Results. A short-flight-test program has been completed at Cornell and some pilot opinions obtained on the relative merits of various configurations. Some of the findings are as follows:

1 Phugoid damping is very evident, but more appreciated by amateur than by professional pilots.

2 Dutch-roll damping by $C_{n\dot{\beta}}$ makes airplane more sensitive to gusts than $-C_{n_r}$. $C_{n\dot{\beta}}$ requires less rudder co-ordination in steady turns.

3 Reduction of roll inertia by $C_{n_{\dot{p}}}$ was liked as it not only improved Dutch-roll damping, but improved rolling response.

4 Spiral stability provided by $-C_{l_r}$ was liked by all pilots, especially on instrument flight. Large amounts of $-C_{l_r}$ require holding aileron deflection during steady turns like steering an automobile. This was liked better by amateur than professional pilots.

CONCLUSIONS

The C-45 airplane just described is one of several artificial stability aircraft which have been flown or are now flying at Cornell With them, by giving pilots the proper tasks to try in flight, we are learning what makes a good airplane from the point of view of both static and dynamic stability and control. Fig. 16 shows the military requirements for acceptable period and damping of the lateral motion as they existed in 1948, before they were placed under security wraps. These requirements represent minimums and while we are interested in minimum acceptable stability we also are concerned greatly with optimum characteristics now that powerful means are available for altering stability through automatic controls.

At the same time, we are learning how to design the automatic control systems which will be necessary to make some of the future airplanes flyable and useful vehicles.

Appendix

NOMENCLATURE

The following nomenclature is used in the Appendix:

General Note. Many stability notations are in use, representing

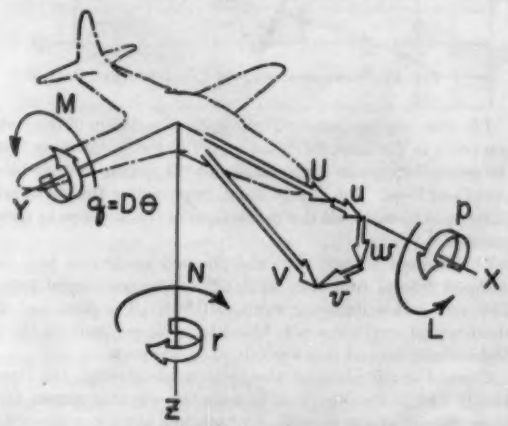


FIG. 17 AXIS SYSTEM

different schemes for nondimensionalizing. Make sure of definition of any quantity before using it literally, or using its numerical value. Different writers may use the same notation with different definitions.

Quantities are positive as shown in Fig. 17.

Stability axes—fixed in and move with aircraft.

X, Y, Z —see Fig. 17 for axis orientation

Velocity components along axes:

U_0 = initial velocity along X, fps
u = incremental velocity along X, fps
v = incremental velocity along Y, fps
w = incremental velocity along Z, fps

Angular velocities around axes:

p = angular velocity around X, rad/sec
$q = D\theta$ = angular velocity around Y,[1] rad/sec
r = angular velocity around Z, rad/sec

Forces along axes:

F_X = force along X, lb
F_Y = force along Y, lb
F_Z = force along Z, lb
T = thrust force along thrust axis, lb

Moments around axes:

L = moment around X, ft-lb
M = moment around Y, ft-lb
N = moment around Z, ft-lb

Aerodynamic angles between X-axis and relative wind (see Fig. 17):

$\alpha = \left(\frac{w}{U_0}\right)$ angle of attack, radians

$\beta = \left(\frac{v}{U_0}\right)$ angle of sideslip, radians

Angular rotation of airplane from equilibrium:

$\phi = \int p\,dt$ bank-angle change around X, radians
$\theta = \int q\,dt$ pitch-angle change around Y, radians
$\psi = \int r\,dt$ heading-angle change around Z, radians

Angular rotation of aerodynamic control surface:

δ_e = elevator deflection (trailing edge down is positive), radians
δ_a = aileron deflection (sum of left + right) (left aileron trailing edge down is positive), radians
δ_r = rudder deflection (trailing edge left is positive), radians

Other angles:

η_T = angle between thrust axis and X-axis (positive when thrust axis is nose-up relative to X), radians
η_P = angle between principal axes and stability axes (positive when nose-up relative to X), radians

Aerodynamic force and moment quantities:

ρ = density of air, slugs per cu ft
S = wing area, sq ft
c = mean aerodynamic chord of wing, ft
b = wing span, ft
q = $\frac{1}{2}\rho U_0^2$—dynamic pressure, psf
M = Mach number, dimensionless
L = aerodynamic lift force in X-Z plane at right angles to velocity vector V, lb
D = aerodynamic drag force in X-Z plane parallel to velocity vector V, lb
F_Y = aerodynamic side force along Y, lb

$C_L = \frac{L}{Sq}$ lift coefficient, dimensionless

$C_D = \frac{D}{Sq}$ drag coefficient, dimensionless

$C_Y = \frac{F_Y}{Sq}$ side-force coefficient, dimensionless

$C_T = \frac{\partial T}{\rho S U_0}$ thrust coefficient, dimensionless

$C_l = \frac{L}{Sqb}$ rolling-moment coefficient, dimensionless

$C_m = \frac{M}{Sqc}$ pitching-moment coefficient, dimensionless

$C_n = \frac{N}{Sqb}$ yawing-moment coefficient, dimensionless

$C_{mT} = \frac{2\frac{\partial M_T}{\partial u}}{\rho S U_0} = \left[2C_{m_0} - 2C_T\frac{Z_T}{c}\right]$ thrust-moment coefficient, dimensionless

where Z_T = height of thrust axis above center of gravity, ft

Stability derivatives in order of their introduction into equations:

Drag derivatives:

$C_{D\mathsf{M}} = \frac{\partial C_D}{\partial \mathsf{M}}$ where M = Mach number, dimensionless

$C_{D\alpha} = \frac{\partial C_D}{\partial \alpha}$

Lift derivatives:

$C_{L\mathsf{M}} = \frac{\partial C_L}{\partial \mathsf{M}}$ where M = Mach number, dimensionless

$C_{L\alpha} = \frac{\partial C_L}{\partial \alpha}$, dimensionless

$C_{L\delta_e} = \frac{\partial C_L}{\partial \delta_e}$, dimensionless

Pitching-moment derivatives:

$C_{m_0} = \frac{\partial C_m}{\partial\left(\frac{u}{U_0}\right)}$, dimensionless

$^*C_{m\dot{u}} = \frac{\partial C_m}{\partial\left(\frac{\dot{u}}{U_0}\right)}$, sec

$C_{m\mathsf{M}} = \frac{\partial C_m}{\partial \mathsf{M}}$, where M = Mach number, dimensionless

$C_{m\alpha} = \frac{\partial C_m}{\partial \alpha}$

$C_{m\dot{\alpha}} = \frac{\partial C_m}{\partial \dot{\alpha}}$, sec

$C_{m\dot{\theta}} = \frac{\partial C_m}{\partial \dot{\theta}}$, sec

[1] q is also used for dynamic pressure. In this paper, the angular velocity is always $D\theta$, and dynamic pressure is always q.

* These derivatives are entirely artificial and do not exist in uncontrolled airplane.

$C_{m\delta_e} = \frac{\partial C_m}{\partial \delta_e}$, dimensionless

$C_{m\dot{\delta}_e} = \frac{\partial C_m}{\partial \dot{\delta}_e}$, sec

Side-force derivatives:

$C_{Y\beta} = \frac{\partial C_Y}{\partial \beta}$, dimensionless

$C_{Y_p} = \frac{\partial C_Y}{\partial p}$, sec

$C_{Y\delta_r} = \frac{\partial C_Y}{\partial \delta_r}$, dimensionless

Rolling-moment derivatives:

$^*C_{l\dot{\beta}} = \frac{\partial C_l}{\partial \dot{\beta}}$, sec

$C_{l\beta} = \frac{\partial C_l}{\partial \beta}$, dimensionless

$^*C_{l\dot{r}} = \frac{\partial C_l}{\partial \dot{r}}$, sec^2

$C_{l_r} = \frac{\partial C_l}{\partial r}$, sec

$^*C_{l\dot{p}} = \frac{\partial C_l}{\partial \dot{p}}$, sec^2

$C_{l_p} = \frac{\partial C_l}{\partial p}$, sec

$C_{l\delta_a} = \frac{\partial C_l}{\partial \delta_a}$, dimensionless

$C_{l\delta_r} = \frac{\partial C_l}{\partial \delta_r}$, dimensionless

Yawing-moment derivatives:

$^*C_{n\dot{\beta}} = \frac{\partial C_n}{\partial \dot{\beta}}$, sec

$C_{n\beta} = \frac{\partial C_n}{\partial \beta}$, dimensionless

$^*C_{n\dot{r}} = \frac{\partial C_n}{\partial \dot{r}}$, sec^2

$^*C_{n_p} = \frac{\partial C_n}{\partial p}$, sec

$C_{n\delta_r} = \frac{\partial C_n}{\partial \delta_r}$, dimensionless

$C_{n\delta_a} = \frac{\partial C_n}{\partial \delta_a}$, dimensionless

Inertia quantities:

m = airplane mass, slugs

W = mg—airplane weight, lb

$I_{XX} = \int (y^2 + z^2) dm$—moment of inertia about X-axis, slugs-ft^2

$I_{YY} = \int (x^2 + z^2) dm$—moment of inertia about Y-axis, slugs-ft^2

$I_{ZZ} = \int (x^2 + y^2) dm$—moment of inertia about Z-axis, slugs-ft^2

$I_{XZ} = \int xz dm$—product of inertia with respect to X and Z-axes, slugs-ft^2

Axes X, Y, Z are nearly coincident with the principal axes of inertia of the airplane for which the moments of inertia are I_{XP}, I_{YP}, I_{ZP}.

By definition, product of inertia for principal axes vanishes, $I_{XZ(P)} = 0$

$$I_{XX} = I_{XP} \cos^2 \eta_P + I_{ZP} \sin^2 \eta_P$$
$$I_{YY} = I_{YP}$$
$$I_{ZZ} = I_{ZP} \cos^2 \eta_P + I_{XP} \sin^2 \eta_P$$
$$I_{XZ} = -(I_{ZP} - I_{XP}) \sin \eta_P \cos \eta_P$$

or for small values of η_P, approximately

$$I_{XX} \approx I_{XP}$$
$$I_{YY} \approx I_{YP}$$
$$I_{ZZ} \approx I_{ZP}$$
$$I_{XZ} \approx (I_{XP} - I_{ZP})\eta_P$$

Miscellaneous symbols and operators:

t = time, sec

ω = frequency of oscillation, rad/sec

ω_n = undamped natural frequency, rad/sec

ξ = critical damping ratio, dimensionless

$D \equiv \frac{d}{dt}$, 1/sec

Dot over variable indicates time derivative

$\dot{\theta} \equiv D\theta \equiv \frac{d\theta}{dt}$, 1/sec

$j = \sqrt{-1}$, dimensionless

g = acceleration of gravity, ft/sec^2

a_Z = acceleration along Z-axis, ft/sec^2

$n_Z = \frac{a_Z}{g}$—acceleration along Z-axis, dimensionless

Derivation of Equations

Fig. 17 shows an airplane and three orthogonal axes fixed in the structure. The airplane and axes have an instantaneous velocity V with components $(U_0 + u)$, v, and w in the X, Y, and Z-directions, and an angular velocity Ω with components p, $D\theta$, and r about the X, Y, and Z-axes, respectively.

The equations of motion are obtained by equating the aerodynamic and gravity forces and moments to the rates of change of linear and angular momenta. These equations are as follows

$$mU_0\left[\left(\frac{\dot{u}}{U_0}\right) + \left(\frac{w}{U_0}\right) D\theta - \left(\frac{v}{U_0}\right) r\right] = \Sigma F_X \ldots\ldots [1]$$

$$mU_0\left[\left(\frac{\dot{v}}{U_0}\right) + \left(\frac{U_0 + u}{U_0}\right) r - \left(\frac{w}{U_0}\right) p\right] = \Sigma F_Y \ldots [2]$$

$$mU_0\left[\left(\frac{\dot{w}}{U_0}\right) + \left(\frac{v}{U_0}\right) p - \left(\frac{U_0 + u}{U_0}\right) D\theta\right] = \Sigma F_Z \ldots [3]$$

$$I_{XX}\dot{p} + (I_{ZZ} - I_{YY}) r D\theta - I_{XZ}(\dot{r} + pD\theta) = \Sigma L \ldots [4]$$

$$I_{YY} D^2\theta + (I_{XX} - I_{ZZ}) rp - I_{XZ}(r^2 + p^2) = \Sigma M \ldots [5]$$

$$I_{ZZ}\dot{r} + (I_{YY} - I_{XX}) pD\theta - I_{XZ}(\dot{p} - rD\theta) = \Sigma N \ldots [6]$$

For small values of the perturbations u, v, w, and p, $D\theta$, and r, it is possible to linearize Equations [1] through [6] by neglecting products of the variables to give

$$mU_0\left\{\left(\frac{\dot{u}}{U_0}\right)\right\} = \Sigma F_X \ldots\ldots\ldots\ldots [7]$$

$$mU_0\left\{\left(\frac{\dot{v}}{U_0}\right)\right\} = \Sigma F_Y \quad [8]$$

$$mU_0\left\{\left(\frac{\dot{w}}{U_0}\right)\right\} = \Sigma F_Z \quad [9]$$

$$I_{XX}\dot{p} - I_{XZ}\dot{r} = \Sigma L \quad [10]$$

$$I_{YY}D^2\theta = \Sigma M \quad [11]$$

$$I_{ZZ}\dot{r} - I_{XZ}\dot{p} = \Sigma N \quad [12]$$

The force and moment summations are made and expressed in stability-derivative notation explained in the Nomenclature.[8] When this is done, Equations [7], [9], and [11] become the longitudinal equations of motion [13], [14], and [15], while [8], [10], and [12] become the lateral equations of motion

$$[mU_0D + (Sq)(2C_D - 2C_T + \mathbf{M}C_{D\mathbf{M}})]\left(\frac{u}{U_0}\right) + [(Sq)(C_{D\alpha} - C_{L\delta})](\alpha) + [W](\theta) = 0 \quad [13]$$

$$[(Sq)(2C_L + 2\eta_T C_T + \mathbf{M}C_{L\mathbf{M}})]\left(\frac{u}{U_0}\right) + [mU_0D + (Sq)(C_{L\alpha} + C_D)](\alpha) + [-mU_0D](\theta) = [(-Sq)C_{L\delta}](\delta_e) \quad [14]$$

$$[(-Sqc)(C_{mu} + C_{m\dot{u}}D + C_{mT} + \mathbf{M}_0C_{m\mathbf{M}})]\left(\frac{u}{U_0}\right) + [(-Sqc)(C_{m\alpha} + C_{m\dot{\alpha}}D)](\alpha) + [I_YD^2 - (Sqc)C_{m\dot{\theta}}D](\theta) = [(Sqc)(C_{m\delta} + C_{m\dot{\delta}}D)](\delta_e) \quad [15]$$

Long-Period Longitudinal (Phugoid) Oscillation

The phugoid oscillation may be analyzed by considering the lift and drag force, Equations [13] and [14], and neglecting changes in angle of attack and the pitching moment equation. If we further simplify by dropping the Mach number terms we can write Equations [13] and [14] with an incremental lift term ΔL as a forcing function in the lift-force equation

$$[mU_0D + 2Sq(C_D - C_T)]\left(\frac{u}{U_0}\right) + [W](\theta) = 0 \quad [16]$$

$$[2SqC_L]\left(\frac{u}{U_0}\right) + [-mU_0D](\theta) = (\Delta L) \quad [17]$$

Solving for speed change

$$\frac{\left(\frac{u}{U_0}\right)}{(\Delta L)} = \frac{\begin{vmatrix}[0] & [W]\\ [1] & [-mU_0D]\end{vmatrix}}{\begin{vmatrix}[mU_0D + 2Sq(C_D - C_T)] & [W]\\ [2SqC_L] & [-mU_0D]\end{vmatrix}} \quad [18]$$

Evaluating the determinants and substituting $(j\omega)$ for D

$$\frac{\left(\frac{u}{U_0}\right)}{(\Delta L)}(j\omega) = \left[\frac{1}{2W}\right]\frac{1}{1 - \left(\frac{\omega}{\omega_n}\right)^2 + j2\xi\left(\frac{\omega}{\omega_n}\right)} \quad [19]$$

where

$$\omega_n = \frac{1}{U_0}g\sqrt{2} = \frac{45.5}{U_0}\ (\text{rad/sec}) \quad [20]$$

$$\xi = \frac{1}{\sqrt{2}}\frac{(C_D - C_T)}{C_L} \quad [21]$$

[8] An excellent treatment of all phases of stability-derivative development is given in BuAer Report AE-61-4 "Fundamentals of Design of Piloted Aircraft Flight Control Systems," vol. 2—"Dynamics of the Air Frame."

Equation [20] shows that the phugoid natural frequency varies inversely with the true speed of the airplane, while Equation [21] indicates that the damping ratio of the motion depends on the difference between the drag and thrust coefficients in relation to the lift coefficient. A low-drag airplane with jet-power plants usually will have a particularly low phugoid damping ratio.

Longitudinal Short-Period Oscillation

Considering the lift and pitching moment Equations [14] and [15], and considering air speed constant at the frequencies of the short-period mode we, again, can write two simplified equations without Mach number and other secondary effects

$$[mU_0D + Sq(C_{L\alpha} + C_D)](\alpha) + [-mU_0D](\theta) = [-SqC_{L\delta}](\delta_e) \quad [22]$$

$$[-Sqc(C_{m\alpha} + C_{m\dot{\alpha}}D)](\alpha) + [I_YD^2 - SqcC_{m\dot{\theta}}D](\theta) = [Sqc(C_{m\delta} + C_{m\dot{\delta}}D)](\delta_e) \quad [23]$$

Solving for α we get

$$\frac{(\alpha)}{(\delta_e)} = \frac{\begin{vmatrix}[-SqC_{L\delta}] & [-mU_0D]\\ [Sqc(C_{m\delta} + C_{m\dot{\delta}}D)] & [I_YD^2 - SqcC_{m\dot{\theta}}D]\end{vmatrix}}{\begin{vmatrix}[mU_0D + Sq(C_{L\alpha} + C_D)] & [-mU_0D]\\ [-Sqc(C_{m\alpha} + C_{m\dot{\alpha}}D)] & [I_YD^2 - SqcC_{m\dot{\theta}}D]\end{vmatrix}} \quad [24]$$

Evaluating the determinants and substituting $(j\omega)$ for D

$$\left(\frac{\alpha}{\delta_e}\right)(j\omega) = [K]\left\{\frac{1 + j\omega\tau}{1 - \left(\frac{\omega}{\omega_n}\right)^2 + j2\xi\left(\frac{\omega}{\omega_n}\right)}\right\} \quad [25]$$

$$K \equiv \left(-\frac{C_{m\dot{\delta}}}{C_{m\alpha}}\right)\frac{\left[1 + \frac{SqC_{L\dot{\delta}}C_{m\dot{\theta}}}{mU_0C_{m\dot{\delta}}}\right]}{\left[1 + \frac{Sq(C_{L\dot{\delta}} + C_D)C_{m\dot{\theta}}}{mU_0C_{m\alpha}}\right]} \quad [26]$$

$$\tau \equiv \left(\frac{C_{m\dot{\delta}}}{C_{m\dot{\delta}}}\right)\frac{\left[1 - \frac{I_YC_{L\dot{\delta}}}{mU_0cC_{m\dot{\delta}}}\right]}{\left[1 + \frac{SqC_{L\dot{\delta}}C_{m\dot{\theta}}}{mU_0}\right]} \quad [27]$$

$$\omega_n \equiv \sqrt{\frac{-SqC_{m\alpha}}{I_Y}}\sqrt{1 + \frac{Sq(C_{L\alpha} + C_D)C_{m\dot{\theta}}}{mU_0C_{m\alpha}}} \quad [28]$$

$$\xi \equiv (-C_{m\dot{\theta}} - C_{m\dot{\alpha}})\sqrt{\frac{Sqc}{I_Y}}\frac{\left[1 - \frac{I_Y(C_{L\alpha} + C_D)}{mU_0c(C_{m\dot{\theta}} + C_{m\dot{\alpha}})}\right]}{\sqrt{1 + \frac{Sq(C_{L\alpha} + C_D)C_{m\dot{\theta}}}{mU_0C_{m\alpha}}}} \quad [29]$$

Lateral Equations of Motion

The three lateral equations of motion are presented in the following for reference. The terms in Equations [31] and [32] are dimensional moments while the terms in Equation [30] represent the first-time derivative of dimensional forces

$$[mU_0D^2 - SqC_{Y\beta}D](\beta) + [mU_0D](r) + [-W - SqC_{Y_p}D](p) = [SqC_{Y\delta_r}D](\delta_r) \quad [30]$$

$$[(-Sqb)(C_{l\dot{\beta}}D + C_{l\beta})](\beta) + [-I_{xz}D - (Sqb)(C_{l\dot{r}}D + C_{lr})](r) + [I_{XX}D - (Sqb)(C_{l\dot{p}}D + C_{lp})](p) = [SqbC_{l\delta_a}](\delta_a) + [SqbC_{l\delta_r}](\delta_r) \quad [31]$$

$$[(-Sqb)(C_{n\dot{\beta}}D + C_{n\beta})](\beta) + [I_{ZZ}D - (Sqb)(C_{n\dot{r}}D + C_{nr})](r) + [-I_{XZ} - (Sqb)(C_{n\dot{p}}D + C_{np})](p) = [SqbC_{n\delta_r}](\delta_r) + [SqbC_{n\delta_a}](\delta_a) \quad [32]$$

Discussion

H. P. Whitaker.[4] The author has presented a very good review of the factors affecting the stability and control of aircraft, and of the increasing use of automatic equipment in modern aircraft. The Cornell Aeronautical Laboratory has made very significant contributions in this field and it is a pleasure for the writer to have this opportunity of recognizing their achievements.

The concept of using a stability computer to change the dynamic performance function of an aircraft is becoming more and more commonplace each year. The ability to modify the inherent fixed control-surface stability through the actuation of the control surface in response to signals arising from suitable sensing devices has made it possible for the aircraft designer to cope with the aircraft stability and control problems arising from the continuing increase in the difference between maximum speed and landing speed. Many aircraft designers today plan to use such devices from the start of the design proposal and expect that they will be used continuously. The Dutch-roll damper installed in the B-47 was one of the first operational aircraft to use this concept. In time it is expected that all high-speed aircraft will be equipped with similar electromechanical devices to achieve desired characteristics of the aircraft about each of its three axes.

In the early days of aviation, Langley and other pioneers attempted to build into their aircraft complete aerodynamic stability so that the pilot had only to steer the aircraft. The Wright brothers showed that this was not necessary, or even desirable, and they achieved success by providing the pilot with suitable control surfaces that would enable him to manuever and to provide the needed stability. In a completely analogous manner, the addition of stability computers relieves the aerodynamic designer from the necessity of providing inherent aerodynamic stability throughout the Mach number range of the aircraft.

We have now reached the limitations of the ability of the pilot to achieve complete stability throughout the flight regime. And so, just as the addition of movable control surfaces relieved the early designer of the necessity of making an inherently stable machine, the addition of stability computers relieves the designer of the necessity of making the pilot-plane combination inherently stable throughout the Mach number range of the aircraft. For this reason the terms—artificial stability, artificial feel, and so forth—are misnomers, and in time the gyros, accelerometers, and other sensing devices together with the servo systems will be considered as completely natural and essential parts of the coordinated aircraft. Fears as to the safety of the aircraft should these devices fail will be overcome as the reliability of the units continually improves.

As indicated by the author, the main customer who has to be satisfied with regard to dynamic behavior of the aircraft is the pilot. The human being is notoriously difficult to analyze rigorously and so the whole problem of pilot likes and dislikes is a very great undertaking in itself. Here again, the Cornell Aeronautical Laboratory has been a leader and has conducted very valuable research. If anything, the paper errs on the side of being a much too modest description of the activities being carried on at that laboratory.

---

[4] Assistant Director, Instrumentation Laboratory, Massachusetts Institute of Technology, Cambridge, Mass.

Disturbance-Response Feedback—A New Control Concept[1]

By J. B. RESWICK,[2] CAMBRIDGE, MASS.

A new scheme of automatic control called disturbance-response feedback is presented. A special form of this method utilizing a dynamic model of the plant in a feedback path is called plant-model control. The classically difficult-to-control plant having a transfer operator consisting of pure delay (transportation lag or dead time) is shown to be brought under optimum control in an experimental study. Disturbance-response feedback requires for optimization the specification of only one parameter. Therefore it provides a useful reference basis for evaluating other schemes. This is done for some simple systems of control. The plant-model control technique offers a possible basis for realizing a completely self-adjusting automatic controller.

Nomenclature

MANY of the definitions and terms used are due to G. A. Philbrick[3] while others are listed in the Report of the AIEE Subcommittee on Feedback Control Systems.[4] The term "plant" is used to represent a system which is to be controlled. In general, a plant will have inputs, i.e., quantities established external to the plant; and outputs or quantities whose values arise from the result of the system response to its inputs. Often one of the inputs is varied to cause one of the outputs to follow a desired operation. An unregulated plant is shown in Fig. 1. In general, the dynamic relationship between a given pair of input-output variables will differ from the one between another pair. Of particular interest is the relationship connecting the output variable to be controlled to the input variable to be manipulated. The symbols m and c will be used for this particular input-output pair while h symbolizes their dynamic transfer operator. The variable c may be considered to be related to m_1 and m_2 through the transfer operators h_1 and h_2. The transfer operators of the other outputs will not be considered since they are not to be controlled. The output c is indicated in Fig. 1 as the sum of hm, a term which may be called the manipulated response, and $h_1m_1 + h_2m_2$ which may be considered to comprise the disturbed response. The subscripts 1 and 2 are intended to indicate the possibility of more complete generalization. The symbols which appear in Fig. 1 are defined as follows:

m = manipulated variable
m_1, m_2 = disturbances
c = controlled variable
c_d = disturbed response
c_c = controlled response
c_1, c_2 = uncontrolled outputs
h = transfer operator between c_c and m
h_1, h_2 = transfer operators relating disturbed response to system disturbances

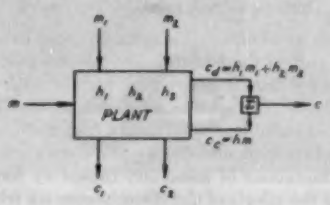


Fig. 1 The Unregulated Plant Showing Defined Variables

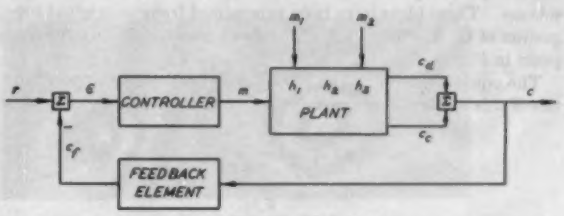


Fig. 2 The Controlled Plant

The objective of automatic control may be considered to have two parts. The first becomes apparent when external disturbances are zero. In this case m must be varied in such a way that c performs in a desired manner. A knowledge of h immediately indicates the proper manipulations of m. This control problem is basically trivial and does not require a closed loop.

When disturbances m_1 and m_2 exist it is clear that m must be varied in a different way to compensate for their additional effects on the output. If the transfer operators h_1 and h_2 are known, and the present and past history of the disturbances as well, it would be possible to define the variation of m to cause c to follow a desired variation. Temperature compensation in a watch is an example where this principle is employed with success. However, it requires a complete knowledge of the transfer operators involved, as well as their variations with time, and in addition a description of the disturbances.

The classical method for solving this problem is to ignore the disturbances on the system and to alter m so as to reduce the difference between a measured value of c called c_f and the value desired for it called the reference variable r. This technique is called automatic regulation or automatic control as contrasted to automatic compensation. Fig. 2 shows the block diagram for this technique. The additional symbols encountered are defined as follows:

---

[1] This paper is based on material which appeared in Chapter V of a thesis entitled "The Design and Application of the Delay-Line Synthesizer," by James B. Reswick. This thesis was submitted to the Mechanical Engineering Department, Massachusetts Institute of Technology, Cambridge, Mass., in June, 1954, as a requirement for the ScD degree.

[2] Assistant Professor of Mechanical Engineering, Massachusetts Institute of Technology. Mem. ASME.

[3] "Unified Symbolism for Regulatory Controls," by G. A. Philbrick, Trans. ASME, vol. 69, 1947.

[4] "Report on Terminology and Nomenclature," by a Subcommittee of the Feedback-Control Systems Committee, AIEE, January, 1950.

Presented at the Instruments and Regulators Division Conference, Ann Arbor, Mich., April 25–26, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, February 4, 1955. Paper No. 55—IRD-2.

Symbols: (not shown in Figure 1)
r = reference variable
ϵ = error; $\epsilon = r - c_f$
c_f = feedback variable

The principle of automatic compensation may be applied to make the task of automatic control easier. Thus, if information concerning the variation of a disturbance is used to cause a corrective effect in m over and above that effect due to the error ϵ alone, a more optimum form of control may be obtained. There are many practical applications of this philosophy; for example, as in the optimization of system controls in the presence of noise disturbances.

It is a specific approach to the problem of automatic control when no information concerning the disturbances is available that is to be presented. In its most general form this novel approach may be identified by the designation "disturbance-response feedback."

Disturbance-Response Feedback

The principle here employed is to make use of information carried in c_e, the part of the output c which is due only to m. This information is fed back positively in parallel with the classical negative feedback path of the total output c. The effects of the manipulated variable m are removed from the operation of the closed loop and consequently the control system is not subject to the classical limitations of instability caused by feedback. Essentially, only the effects of the disturbances are felt around the loop and cause corrective action in the controller; hence the term "disturbance-response feedback" has been chosen for this scheme. These ideas have been generalized from an original suggestion of G. A. Philbrick.[5] A block diagram of this scheme appears in Fig. 3.

The equation defining the error ϵ is indicative of the possibilities of this method of control

$$m = k_1\epsilon + k_2hm = \frac{k_1\epsilon}{1 - k_2h} \quad [1]$$

$$c_e = \frac{k_1h\epsilon}{1 - k_2h} \quad [2]$$

$$\epsilon = r - c = r - c_d - \frac{hk_1\epsilon}{1 - k_2h}$$

$$\epsilon = (r - c_d)\frac{1 - k_2h}{1 + (k_1 - k_2)h} \quad [3]$$

where

$$c_d = h_1m_1 + h_2m_2$$

is the effect of all disturbances. The criterion that the steady-state value of the error ϵ_{ss} be zero, immediately defines the value of k_2 in Equation [3]

$$k_2h_{ss} = 1.0 \quad [4]$$

where h_{ss} is the steady-state limit of h. The characteristic equation of the error response is given by the denominator of Equation [3]. It is this equation which essentially determines the dynamics of the error response. The only parameter which remains to be varied is k_1. Thus the value of k_1 depends on the individual system dynamics, the type of disturbance, and form of criterion used with the disturbance. For example, most systems will have some value of k_1 which will cause the error to be reduced to zero in the minimum possible time without overshoot (aperiodic) in the face of a step change in r. Specific solutions for k_1 in some basic systems are worked out in Appendix 1.

It is interesting to note that for the control of a plant consisting of a simple lag, disturbance-response feedback leads to a proportional-plus-integral controller having a proportional gain of unity and an integrating time constant equal to the time constant of the simple lag. In the case of a second-order system, it leads to a conventional three-term controller, but with a simple lag in series. This indicates that for an aperiodic response, disturbance-response feedback approaches, but is not quite as good as the optimum controller having as many terms (derivatives) as the characteristic equation of the plant to be controlled (see Appendix 1).

The practical considerations in applying this technique as presented in Fig. 3 require further explication. However, this subject

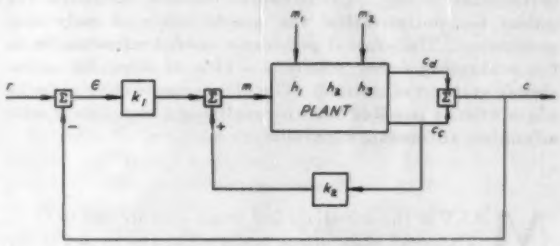


Fig. 3 Disturbance-Response Feedback

will be referred to in a later section. The most significant consideration will be how it may be possible to extract the manipulated response c_e from the total output c. A definite possibility involves the use of a model of the plant to be controlled.

Plant-Model Controller

The plant-model controller (PMC for brevity) is a scheme which utilizes a dynamic model of that part of the plant which relates the manipulated response c_e to the input m. This obviates the need for separating out of the total output, that part which is undisturbed. Fig. 4 illustrates this scheme in block-diagram form.

The real significance of the plant-model controller is that it offers a practical way to utilize and study disturbance-response feedback principles. Two delay-line synthesizer units described in a paper by T. P. Goodman and the present author entitled, "Determination of System Characteristics From Normal Operating Records" permit ready analog evaluation of the scheme. In addition, certain actual plants may be controlled through use of models which are economically feasible.

Analog Evaluation of Plant-Model Controller

Optimum Conventional Control of Pure Delay. One of the classically most difficult plants to control is that which is characterized by the operation of pure delay or dead time. Such a plant has a unit transference

$$h = e^{-TD} \quad [5]$$

It has been shown[6,7] that the error ϵ, due to a step change in r may not be reduced without overshoot to zero in a time which is less than about 3 times the delay time T of the plant by a con-

[5] Discussion with G. A. Philbrick during October, 1952, during which the plant-model control scheme of Fig. 4 was revealed by Mr. Philbrick.

[6] "Time Lag in a Control System," by D. R. Hartree, A. Porter, A. Callander, and A. B. Stevenson, Proceedings of the Royal Society of London, England, series A, vol. 161, 1937, pp. 460–476.

[7] "On the Automatic Control of Generalized Passive Systems, by Kun Li Chien, J. A. Hrones, and J. B. Reswick, Trans. ASME, vol. 74, 1952.

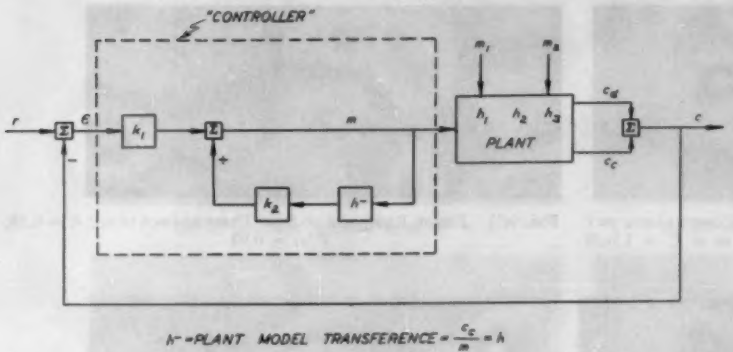


FIG. 4 PLANT-MODEL CONTROL SCHEME

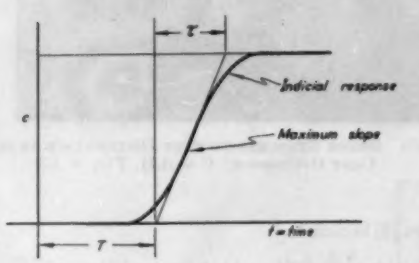


FIG. 5 τ AND T DEFINED. τ/T IS USED TO CHARACTERIZE DYNAMIC RESPONSE OF A PASSIVE SYSTEM

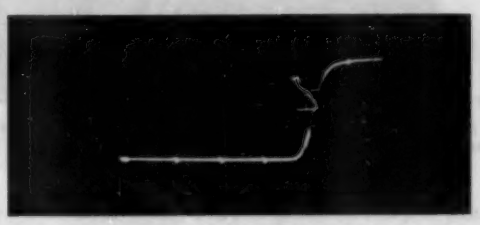


FIG. 6 INDICIAL RESPONSE OF DLS SET FOR UNIT DELAY WITH 30 DELAY UNITS USED
(Time markers are 2.0 millisec apart. Total delay time T = 8.0 millisec.)

FIG. 7(*a*) ERROR RESPONSE ϵ OF LARGE DELAY PLANT UNDER PROPORTIONAL CONTROL FOR r DISTURBANCE. CONTROLLER GAIN C = 0.37

FIG. 7(*b*) ERROR RESPONSE ϵ FOR SAME SYSTEM FOR m_1 DISTURBANCE. CONTROLLER GAIN C = 0.43

troller having proportional, derivative, and integral terms. Chien, Hrones, and Reswick[7] deal with the control of passive plants having significant delay. Such plants were characterized by the ratio of τ/T where τ and T are defined as in Fig. 5. In the Chien, Hrones, and Reswick investigations, the analog equipment used did not permit values of τ/T less than 1 to be obtained. A newly developed device, the delay-line synthesizer,[8] permits a value of τ/T = 0.16 to be obtained. The indicial response of such a plant is shown in Fig. 6.

Since this plant approaches the plant with pure delay, it was studied first to determine the optimum with proportional control only, then with proportional-plus-integral control. This was done in order to extend the data of reference[7] and then to determine the characteristics of PMC when applied to it.

Fig. 7(*a*) shows the error response of the system of Fig. 6 under proportional control subjected to a step disturbance in r while Fig. 7(*b*) shows the same system response for a load disturbance m_1. The criterion is for 20 per cent overshoot as defined in reference.[7]

Fig. 8 shows the error response for the same system with a disturbance with proportional gain set up to incipient instability.

[8] "Determination of System Characteristics From Normal Operating Records," by T. P. Goodman and J. B. Reswick, ASME Paper No. 55—IRD-1.

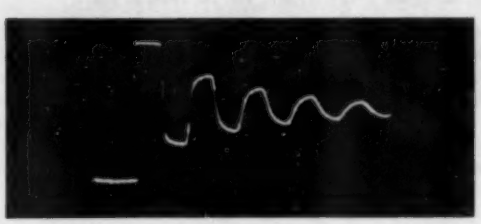


FIG. 8 ERROR RESPONSE ϵ OF LARGE DELAY PLANT UNDER PROPORTIONAL CONTROL FOR r DISTURBANCE. CONTROLLER GAIN C = 1.0

Fig. 9(*a*) shows the error response due to an r disturbance when under proportional-plus-integral control. Parameters have been set for quickest reduction of error to zero with no overshoot. Fig. 9(*b*) shows the error response to a load disturbance under the same conditions. Figs. 10(*a*) and 10(*b*) show the same responses with parameters adjusted for 20 per cent overshoot.

The results given in Figs. 8, 9, and 10 may be used to extend the results plotted by Chien, Hrones, and Reswick.[7] Some pertinent graphs have been replotted with the new data in Appendix 2. It is interesting to note that the value of loop gain of unity which

FIG. 9(a) APERIODIC ERROR RESPONSE TO STEP DISTURBANCE IN r WITH PROPORTIONAL-PLUS-INTEGRAL CONTROL; $m = C + 1/\tau_i D$ $C = 0.25$, $T/\tau_i = 0.90$

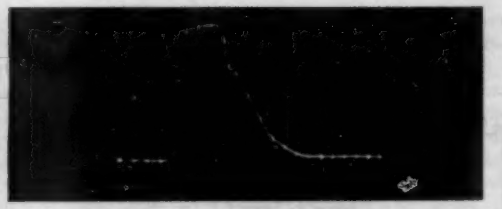


FIG. 9(b) ERROR RESPONSE TO STEP DISTURBANCE IN m; $C = 0.28$, $T/\tau_i = 0.92$

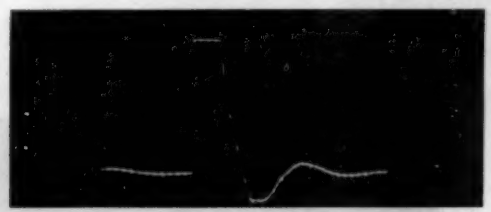


FIG. 10(a) ERROR RESPONSE TO STEP DISTURBANCE IN r WITH PROPORTIONAL-PLUS-INTEGRAL CONTROL; $m = C + 1/\tau_i D$ 20 PER CENT OVERSHOOT; $C = 0.41$, $T/\tau_i = 1.3$

FIG. 10(b) ERROR RESPONSE TO STEP DISTURBANCE IN m; 20 PER CENT OVERSHOOT; $C = 0.41$, $T/\tau_i = 1.3$

produced incipient instability in Fig. 8 (proportional control of a plant with large delay) is shown by Oldenbourg and Sartorius[9] to be the theoretical value for the control of a pure delay plant.

PMC of Pure Delay. In studying the control of a plant having pure delay using PMC, a somewhat shorter delay time than that pictured in Fig. 6 was chosen. This permitted one DLS unit to be set up as a plant and a second DLS unit set up as a model. Each unit was set up to have the indicial response shown in Fig. 12(*a*). The DLS units were then connected in the form of the block diagram of Fig. 4. In accordance with Equation [4], k_2 was set at unity. That k_1 should also be set at unity is apparent from the following:

If

$$h = e^{-TD} = h \text{ and } k_1 = k_2 = 1$$

[9] "The Dynamics of Automatic Controls," by R. C. Oldenbourg and H. Sartorius, Trans. ASME, vol. 70, 1948, p. 104.

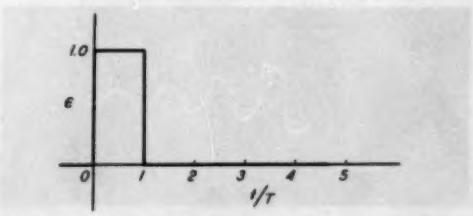


FIG. 11 THEORETICAL ERROR RESPONSE FOR PLANT WITH PURE DELAY UNDER PMC; $k_1 = k_2 = 1$

Equation [3] becomes

$$\epsilon = (r - c_d)(1 - h) = (r - c_d)(1 - e^{-TD})$$

Thus for a step change in r, $c_d \equiv 0$, the error will appear as in Fig. 11.

The actual error resulting from a step change in r is shown in Fig. 12(*b*). Note that the plant indicial response is mirrored in the error response with the result that the error is brought to zero within $\frac{1}{2}$ of one delay time. The improvement over the performance of Fig. 9(*a*) is immediately apparent.

In an attempt to determine the effect of variations in k_1 and k_2, a series of four tests were made where in each case either k_1 or k_2 was increased or decreased by 25 per cent. In the case of a variation in k_1, one would expect a change in the dynamic portion of the response, but with the steady-state error remaining at zero. In the case of a variation in k_2, a change both in the steady-state error (25 per cent) and a change in the dynamics would be expected. These effects are indeed shown in Figs. 13(*a*, *b*, *c*, *d*).

Addition of Integral Term to PMC for Reduction of Steady-State Error. As has been shown, a variation in k_1 which represents a change in over-all loop gain arising either in the plant or controller changes the dynamics of the response as it would in a conventional controller, but it does not affect the steady-state error. However, a change in k_2 which causes the loop gain of the model controller to depart from unity is followed by a change in the steady-state error.

Since the gain of the positive loop in the model controller is determined only by equipment within the control mechanism, many

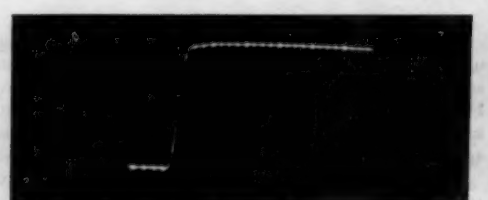


FIG. 12(a) INDICIAL RESPONSE OF PLANT h AND OF PLANT MODEL $\hat{h}$

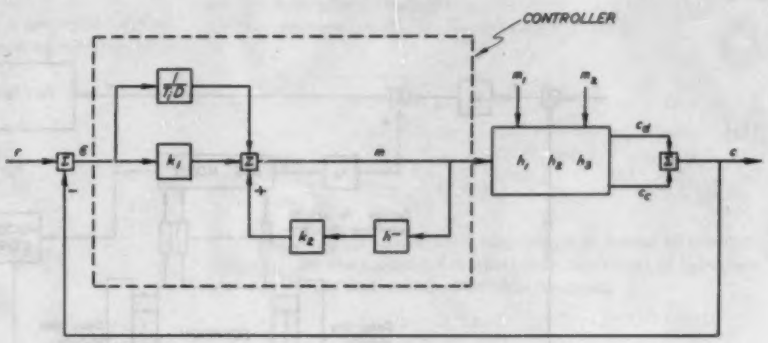
FIG. 12(b) ERROR RESPONSE OF PLANT-MODEL CONTROL SCHEME FOR STEP CHANGE IN r

(*a*) 25 per cent decrease in k_1, $k_2 = 1$ (*b*) 25 per cent increase in k_1, $k_2 = 1$

(*c*) 25 per cent decrease in k_2, $k_1 = 1$ (*d*) 25 per cent increase in k_2, $k_1 = 1$

FIG. 13 EFFECT OF VARIATION OF SYSTEM PARAMETERS ON ERROR RESPONSE FOR OPTIMUM PMC OF SYSTEM WITH PURE DELAY

FIG. 14 PLANT-MODEL CONTROL SCHEME WITH ADDITION OF INTEGRAL TERM



physical possibilities exist where changes in gain may be prevented from occurring. The possibility of compensating for such changes through the addition of an integral term exists and some experimental evaluation of one scheme follows.

Fig. 14 shows a block diagram which is similar to that of Fig. 4 but with the addition of an integral effect. Fig. 15 shows the error responses due to an r disturbance with and without the integral term for cases where k_2 is greater and less than unity. The integral time constant was adjusted for an arbitrary optimum. It is interesting to note that this optimum gives the same result for T_i as the optimum choice for conventional proportional-plus-integral control gives, see Fig. 10.

As is evident in the results, the operation of the plant-model controller under conditions where the gain of the positive feedback loop of the controller deviates from unity by 25 per cent is inferior to the operation of an optimumly adjusted proportional-plus-reset controller, even with the addition of an added integral. It is felt, however, that a 25 per cent change in gain of this loop represents a very drastic basis of comparison. In cases where this gain is mechanically determined, added integral effects are unnecessary and the over-all performance is quite superior. Where the gain may vary slightly, the addition of the integral term would insure perfect reduction of steady-state error, but still retaining the improvement in performance. If the actual design must be done in a way that does not permit fair stability in gain of the PMC loop, it appears that the PMC should be rejected in favor of a more conventional control system.

AUTOMATIC SELF-ADJUSTING CONTROLLER

The plant-model controller is limited in its ability to control by the departures of the model transference from the transference of the undisturbed plant. In some cases this transference may be very difficult to determine, and in addition, it usually will not remain constant with time, or even with the particular operating point of the system. The existence of high-speed correlating equipment such as designed by Bell[10] suggests the possibility of continuously determining the undisturbed transfer operator of a plant through auto and cross correlation of c with m. Fig. 16 shows a scheme which would calculate continuously from the auto and cross-correlation functions, the delay series representation of the controlled-response weighting function. These values then would be used to command servos which would set the knobs of a DLS to become a correct linear model based on an analysis of a section of past data. Even though the plant might be nonlinear, the fact that at any instant it is controlled by an optimum controller, based on the linear equivalent plant for a short period of past history, may result in an optimum over-all control stability.

DISTURBANCE-RESPONSE FEEDBACK—POSSIBILITIES OF REALIZATION

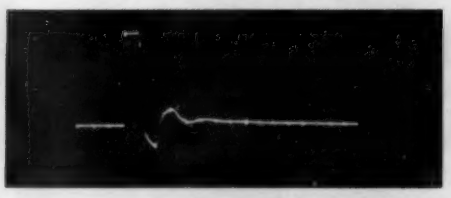
The basic problem in utilizing the pure scheme of disturbance-response feedback as suggested in Fig. 3, involves the determina-

[10] "The Design and Some Applications of a High-Speed Correlator," by H. Bell, Jr., PhD thesis, University of Wisconsin, Madison, Wis. 1953.

(a) Error response to step change in r for 25 per cent increase in k_2

(b) Error response under same conditions with added integral term $T/\tau_i = 0.88$

(c) Error response for step change in r for 25 per cent decrease in k_2



(d) Error response under same conditions with added integral term $T/\tau_i = 0.88$

FIG. 15 ERROR RESPONSE

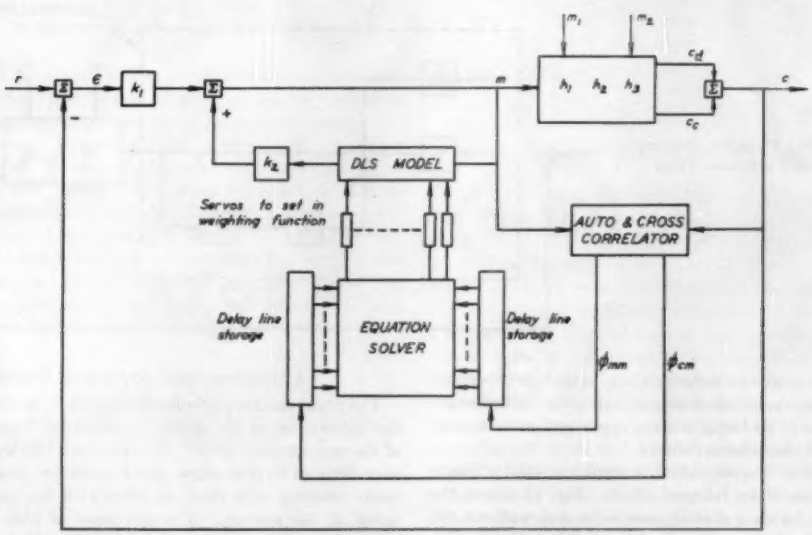


FIG. 16 AUTOMATIC SELF-ADJUSTING CONTROLLER

tion of that part of the output c which is due only to the manipulated input m. The use of a model is clearly a possibility, but the advantages of obtaining c_c directly from c would be tremendous; for in this case the plant would be its own model and must therefore always be under optimum control. Nonlinear industrial processes might be started up quickly and stopped quickly without human interventions.

No real solution to this problem will be offered here, but some suggestions may be clues to actual application. It is possible, for example, that in a chemical process, some variable other than that desired to be controlled (ph, density, color, etc.) may not be subject to the same disturbances or to the same degree as the variable under control. A measurement and feedback of such a variable might approximate a disturbance response-feedback scheme. The possibility of sending a radioactive isotope through a process in such a way that it is affected only by the undisturbed dynamics of the plant, does not seem too farfetched to rule out.

A process utilizing its own dynamics to control itself in an optimum fashion is the tantalizing prospect of these thoughts. A specific limited solution is given in the form of PMC, but a final answer not needing a model presents a goal not to be discarded lightly.

Appendix 1

SOME COMPARISONS OF PMC OF SIMPLE SYSTEMS WITH CORRESPONDING CONVENTIONAL CONTROL

The plant-model control scheme is shown in Fig. 4. The error response is given by Equation [3]

$$\epsilon = (r - h_1 m - h_2 n)\left[\frac{1 - k_2 h}{1 + (k_1 - k_2)h}\right] \quad \ldots\ldots [3]$$

Equation [4] which determines k_2 such that $h_{ss} = 0$ is given by

$$k_2 h_{ss} = 1.0 \quad [4]$$

The transfer operator of the plant-model "controller," see Fig. 4, is given by

$$\frac{m}{\epsilon} = \frac{k_1}{1 - k_2 h} \quad [6]$$

Selection of k_1 depends on the criterion placed on the characteristic equation of the error response. The criterion of aperiodicity (no significant imaginary roots) will be applied to some simple cases in order to define k_1 and to compare the effective plant-model controller with known optimum conventional ones.

Unit Lag Plant

Let

$$h = \frac{1}{1 + \tau D}$$

Controller is (from Equation [6])

$$\frac{m}{\epsilon} = \frac{k_1}{1 - \dfrac{1}{1 + \tau D}} = \frac{k_1(1 + \tau D)}{\tau D} = k_1 \left(1 + \frac{1}{\tau D}\right) \quad [7]$$

The controller equation is seen to be that of a proportional-plus-integral controller in which the time constant of integration is made equal to the time constant of the unit lag.

The error response to a disturbance in r is given by

$$\frac{\epsilon}{r} = \frac{1 - \dfrac{1}{1 + \tau D}}{1 + \dfrac{(k_1 - 1)}{1 + \tau D}} = \frac{\tau D}{\tau D + k_1} \quad [8]$$

Fig. 17 shows typical responses for various k_1 values.

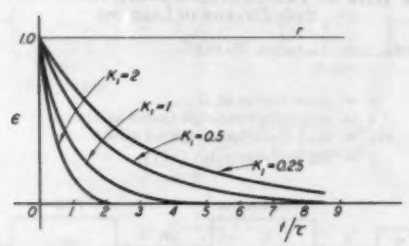


FIG. 17 ERROR RESPONSES TO STEP CHANGE IN r—PMC OF UNIT LAG

Second-Order Plant

Let

$$h = \frac{1}{\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D + 1}$$

Plant-model controller becomes

$$\frac{m}{\epsilon} = \frac{k_1}{1 - \dfrac{1}{\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D + 1}} = \frac{k_1 \left(\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D + 1\right)}{\left(\dfrac{1}{\omega_n^2} D + \dfrac{2\zeta}{\omega_n}\right) D}$$

$$\frac{m}{\epsilon} = \frac{k_1}{\left(\dfrac{1}{2\zeta\omega_n} D + 1\right)} \left(\frac{1}{2\zeta\omega_n} D + 1 + \frac{\omega_n}{2\zeta D}\right) \quad [9]$$

This equation has the form of a proportional-plus-integral-plus-derivative controller with a series lag of time constant equal to the time constant of the differentiating term. This result indicates that the PMC will not do as well as an optimum 3-term controller when controlling a second-order system. One may generalize that it will never do quite as well as a controller having the same order as the plant to be controlled. However, the use of more than three terms in a controller is not often done. The performance of the PMC will then exceed that of the conventional controller in all plants of order greater than four.

Equation [9] may be interpreted to define an optimum ratio of integral to derivative time constant which might have significance where the three adjustments of a conventional controller have been replaced by two controls, i.e., proportional band and integral-derivative setting. This ratio is defined by

$$\frac{T_i}{T_d} = \frac{\dfrac{2\zeta}{\omega_n}}{\dfrac{1}{2\zeta\omega_n}} = 4\zeta^2 \quad [10]$$

The indication here is that a single value of the time-constant ratio will not suffice for all damping ratios.

The error response to an r disturbance is given by

$$\frac{\epsilon}{r} = \frac{1 - \dfrac{1}{\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D + 1}}{1 + \dfrac{k_1 - 1}{\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D + 1}} = \frac{\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D}{\dfrac{1}{\omega_n^2} D^2 + \dfrac{2\zeta}{\omega_n} D + k_1} \quad [11]$$

For the condition of aperiodic response, it is desired to evaluate k_1 so that the denominator (characteristic equation) of Equation [11] has equal, real roots. For this condition

$$\frac{k_1}{\dfrac{1}{\omega_n^2}} = \frac{\dfrac{2\zeta}{\omega_n}}{\dfrac{2}{\omega_n^2}} = \zeta^2 \omega_n^2 \quad [12]$$

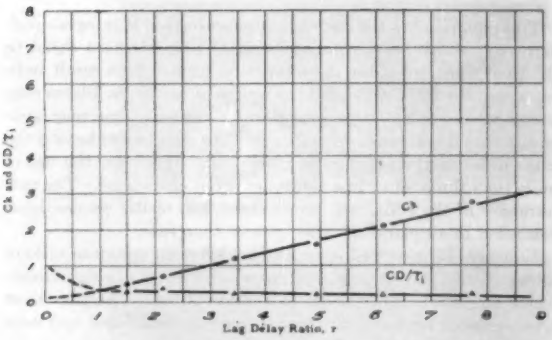
$$k_1 = \zeta^2$$

Higher-Order Plants. The mathematical analysis of higher-order plants (with the exception of those characterized by pure dead time) (see section, Automatic Self-Adjusting Controller) becomes involved. The steps in any given case are straightforward and a value of k_1 may be determined by the criterion which is applied to the characteristic equation of the response of the controlled system. Since k_1 is the only variable parameter in an actual installation it may be determined by experiment.

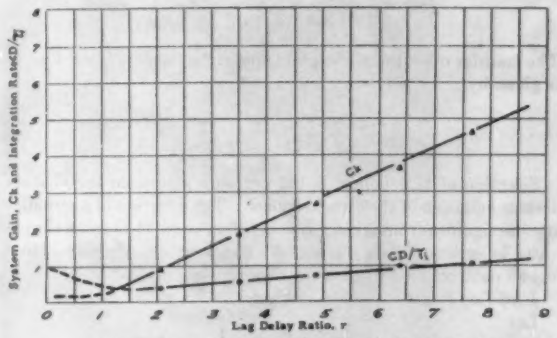
Appendix 2

EXTENSION OF RESULTS OF CHIEN, HRONES, RESWICK[7] FOR PROPORTIONAL-PLUS-INTEGRAL CONTROL OF PASSIVE SYSTEMS

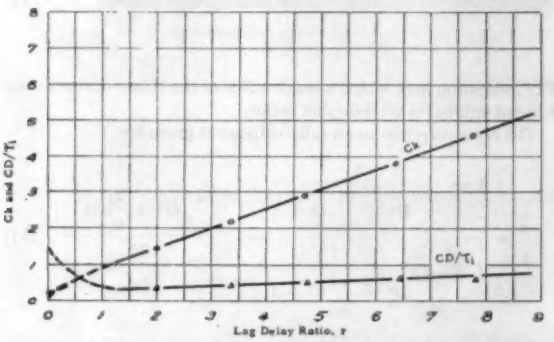
In the reference cited in the heading of this Appendix, optimum controller settings for the control of passive systems defined in terms of $\tau = \tau/T$ as in Fig. 5 are plotted for various disturbances and for various optimum criteria. In Fig. 18 these results have been reproduced with the additional points for $\tau = 0.16$. The dotted portions of the curves have been added to fit the curves together. The added points are in good agreement with the trends indicated in the curves for desired value disturbances, but they differ from the load-disturbance trends in that they are very



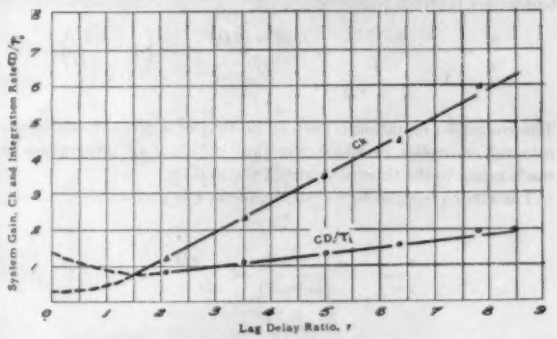
VALUES OF Ck AND CD/τ_i WHICH PRODUCE QUICKEST RESPONSE WITHOUT OVERSHOOT FOR VARIOUS r WITH STEP CHANGE IN DESIRED VALUES, v^*



VALUES OF Ck AND CD/τ_i WHICH PRODUCE QUICKEST RESPONSE WITHOUT OVERSHOOT FOR VARIOUS r WITH STEP CHANGE IN LOAD, m_L



VALUES OF Ck AND CD/τ_i WHICH PRODUCE QUICKEST RESPONSE WITH 20 PER CENT OVERSHOOT FOR VARIOUS r WITH STEP CHANGE IN DESIRED VALUE, v^*



VALUES OF Ck AND CD/τ_i WHICH PRODUCE QUICKEST RESPONSE WITH 20 PER CENT OVERSHOOT FOR VARIOUS r WITH STEP CHANGE IN LOAD, m_L

FIG. 18 EXTENDED RESULTS OF CHIEN, HRONES, AND RESWICK PAPER[7]

FOR $r = 0.16$

Captions on figures apply where
- Ck = loop gain
- D = delay time (same as T)
- v^* = desired value (same as r)
- u = error (same as ϵ)
- v = controlled variable (same as c)
- m_L = load disturbance (same as m_l)
- r = lag in delay ratio (τ/T)

close to the desired value-disturbance results. Apparently the difference in the form of disturbance has decreasing effect on the optimum controller settings as the system approaches a pure delay having a value of $r = 0$.

Discussion

G. LANG[11] AND J. M. HAM.[12] This paper makes a valuable contribution to the techniques for loop compensation in feedback-control systems by providing a particular, simplified procedure for the design and adjustment of a controller for a given plant. The essential element in the scheme is a plant model.

Since we have published papers[13,14] on new forms of control systems employing models, it will be useful to distinguish the various uses for such models. Fig. 19 of this discussion is a simplified form for Fig. 4 of the paper. The box labeled "compensator" is what the author has called a "plant-model controller." The compensator or controller-transfer characteristic

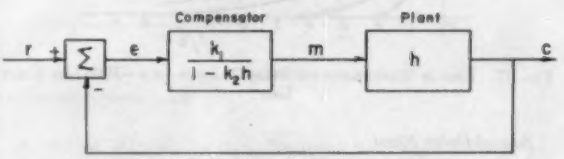


FIG. 19 LOOP COMPENSATION WITH A PLANT MODEL

$$\frac{m}{\epsilon} = \frac{k_1}{1 - k_2 h}$$

is obtained by using a model h of the transfer characteristic of the plant in an internal positive-feedback loop. Adjustment of the compensator is made by setting the gain factors k_1 and k_2. We regard this procedure for compensator design as a valuable addition to the literature on the subject but are rather at a loss to understand why the author calls the scheme "disturbance-response feedback."

[11] Ferranti Electric Ltd., Toronto, Canada.

[12] Department of Electrical Engineering, Universiy of Toronto, Toronto, Canada.

[13] "Conditional Feedback Systems—A New Approach to Feedback Control," by G. Lang and J. M. Ham, Technical Paper 55-202, AIEE, 1955.

[14] "New Concepts in Process Control," by J. M. Ham, *Canadian Chemical Processing*, Toronto, Canada, vol. 39, February, 1955, pp. 72–76.

Consider the contributions of the manipulated variable m and the output disturbance c_d to the loop error signal ϵ. Let these contributions be denoted by ϵ_m and ϵ_{cd}. From Equations [1] and [3] in the paper

$$\frac{\epsilon_m}{\epsilon_{cd}} = \frac{m}{-c_d}\,\frac{1+(k_1-k_2)h}{k_1} \qquad [13]$$

If $k_2 h_{ss} = 1$ as required in the PMC design, Equation [13] of this discussion reduces to

$$\frac{\epsilon_m}{\epsilon_{cd}} = \frac{mh}{-c_d} = \frac{m}{-c_d}\,\frac{1}{k_2} = \frac{c_c}{-c_d} \qquad [14]$$

From Equation [14] herewith, it is clear that ϵ_m is not necessarily small in comparison with ϵ_{cd} so that the transient variations of both m and c_d are felt in the loop just as in the classical feedback-control system. It may be noted that, at zero frequency, both ϵ_m and ϵ_{cd} are vanishingly small because of infinite loop gain. If the system is used as a regulator with r = const, then fluctuations in ϵ are contributed entirely by c_d. But this statement is true for any classical feedback-control system. Hence the use of the designation disturbance-response feedback is questioned.

The author's use of a plant model is distinctly different from our use of models in conditional feedback systems.[13,14] His plant-model controller can be incorporated readily in our conditional feedback systems as a loop-compensating element. This use of his PMC is shown in Fig. 20 of this discussion. Fig. 20 of this discussion represents a conditional feedback-system design corresponding to the author's PMC design for a plant having a pure delay as a transfer characteristic, namely $h = e^{-TD}$. He is to be commended for dealing with the difficult subject of plants with pure delays.

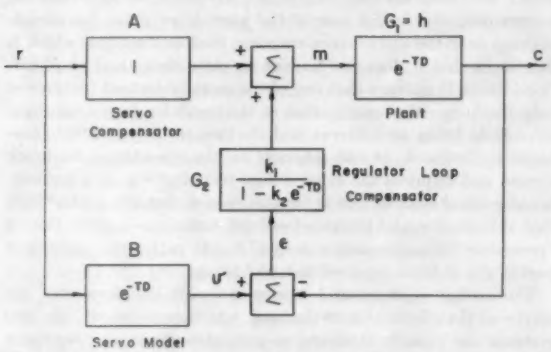


FIG. 20 A CONDITIONAL FEEDBACK SYSTEM EMPLOYING A PLANT-MODEL CONTROLLER AS A LOOP-COMPENSATING ELEMENT

In Fig. 20 of this discussion, the plant is represented by the transfer function $G_1 = h = e^{-TD}$ for this example. A is a cascade compensator for adjusting the servo performance of the system, namely c/r. G_2 is the regulator loop compensator. B is called a "servo model" and the response ratio $v/r = B$ defines the servo performance desired of the plant. This use of a model should be clearly distinguished from that of the author. His model is a model of the transfer characteristic of the actual plant and is used essentially to achieve loop compensation. Our servo or reference model is a model of the servo performance desired of the plant and is used to achieve this performance. The design of our conditional system as shown in Fig. 20 of this discussion proceeds as follows:

If $B = AG_1$, then $v = c$, the signal ϵ in the feedback loop is zero and $c/r = AG_1 = B$. Under these conditions, the servo performance is unaffected by the loop compensator G_2. If there is an additive output disturbance c_d, the loop-error signal $\epsilon = c - (c + c_d) = -c_d$ and is a direct measure of the disturbance. Hence the system of Fig. 20 of this discussion could well be called a disturbance-response feedback system. The condition $B = AG_1$ is called the conditional equation and defines the element A when B and G_1 are given. Since G_2 has no effect on the servo behavior when the conditional equation is satisfied, it can be selected independently to yield the most satisfactory system response to disturbances. For a plant with a pure delay, the best servo performance that can be expected is a pure delay. Hence the system model $B = h = e^{-TD}$ and $A = 1$. To compare the complete conditional design with that of the author, G_2 is taken to be a PMC, namely

$$G_2 = \frac{k_1}{1 - k_2 h} = \frac{k_1}{1 - k_2 e^{-TD}}$$

The response of the conditional system in Fig. 20 of this discussion to an additive output disturbance c_d is identical with that of the author's design in Fig. 4. However, the servo-response ratio c/r of the conditional system is $c/r = e^{-TD}$ independently of k_1 and k_2. Hence the conditional design achieves the best possible servo performance, a simple pure delay. This fact is to be contrasted with the results of Fig. 13 in the paper. The effect of gain changes in the plant itself on the servo response from r to c is the same in both designs. However, the servo response of the conditional design is always independent of the gain of the loop compensator. Such is not the case in the author's PMC design, as his Fig. 13 clearly demonstrates. A further point of contrast between the PMC and conditional system designs is the following: In the PMC design, the forms of the servo and regulator transient response will always be essentially similar, whereas in the conditional design the servo transient response can be designed to be quite different in character from the regulator transient response.

We regard the author's suggestions for controllers that are self-adjusting by means of auto and cross-correlation operations on c and m as very interesting. Proposals of this type recently have been made in electrical communications in connection with "learning filters."[15] The measurement of undisturbed output by means of a reference model is achieved in our conditional feedback systems. The use of "pilot" or "tracer" signals to determine the servo response of systems is commonly used in communication systems and in electronic computing elements[16] and the author's suggestion of the possible significance of this concept in process control is well made.

Models used in PMC designs for loop compensation and in conditional designs as servo-performance references are likely to be significant in solving some of the pressing problems of process control.

AUTHOR'S CLOSURE

The author wishes to express his appreciation to Messrs. Lang and Ham for their detailed discussion. He has been aware of their conditional feedback approach for some time and has felt that similarities exist between this new approach and his disturbance-response feedback. The discussion has clarified many of the points of similarity and dissimilarity in the author's mind.

The words "disturbance-response feedback" have been chosen to designate the author's method, for the following reasons:

[15] "Communication Theory and Cybernetics," by D. Gabor, Trans. IRE, PGCT, vol. CT-1, December, 1954, pp. 19–31.

[16] "Waveforms," by B. Chance, et al., vol. 19, Radiation Laboratory Series, McGraw-Hill Publishing Company, New York, N. Y., 1949, pp. 674–676.

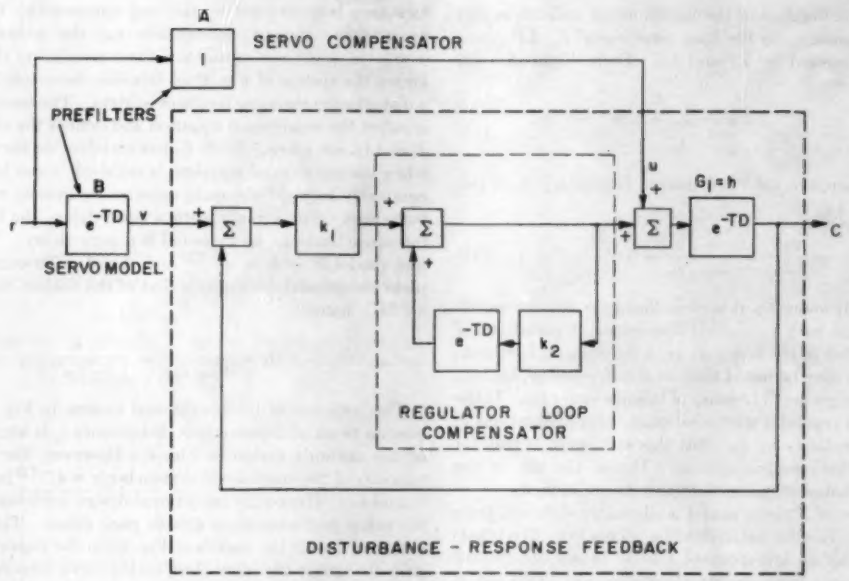


FIG. 21

Consider the system of Fig. 19 with the possibility of an added load disturbance, see Fig. 4.

When there is no load disturbance, the system is described by

$$\frac{c}{r} = \frac{\dfrac{k_1 h}{1 - k_2 h}}{1 + \dfrac{k_1 h}{1 - k_2 h}} = \frac{k_1 h}{1 - k_2 h + k_1 h}$$

Now $k_2 h_{ss} = 1$ and although k_1 is not prescribed in the paper, in general $k_1 h_{ss}$ should be set near unity. To the extent that k_1 is not equal to k_2 these remarks will not apply. When $k_1 = k_2$

$$\frac{c}{r} = k_1 h$$

which is the open loop characteristic of the plant *without feedback*. In effect, the negative feedback of the classical control systems is removed by the positive feedback as far as changes in r are concerned.

Now consider the case when $r = 0$; as Messrs. Lang and Ham point out in their discussion, fluctuations in ϵ and consequently in m are contributed entirely by c_d. This statement is true for any classical feedback-control system as they state, but it is also true that when a change in r occurs in a classical system, fluctuations in m are caused by the output; that this is not true in the disturbance-response feedback scheme has been shown in the foregoing. Since fluctuations in m due to feedback are contributed only by disturbances, it is felt that the designation disturbance-response feedback is particularly appropriate.

The author agrees that conditional feedback systems as described in the discussers' reference are different from disturbance-response feedback systems. In order to further point up the similarities and differences, the author has redrawn Fig. 20 to appear as in Fig. 21.

It appears that the conditional feedback system may be viewed as a conventional control system (system enclosed by heavy dashed line) with inputs u and v where v is similar to the reference or set-point and u a pure load disturbance and in which the *"regulator loop compensator" is undefined*. The author feels the discussers are to be commended for presenting a method for specifying an optimum form and arrangement of prefilters and it is interesting to note that their system inherently requires a plant model (in addition to the regulator loop compensator) to satisfy the condition that $AG_1 = B$. However, he feels that the system proposed in the case of the pure delay plant has no advantage over the disturbance-response feedback scheme which is simpler in that it does not require an extra delay and amplifier. In addition, it appears that one of the most important features of both methods—the specification of the regulator loop compensator—while being an *inherent* and obvious result of disturbance-response feedback, is not inherent to the conditional feedback system and requires the same design techniques as in a conventional control system. It is for this reason that the author feels that the plant-model positive feedback technique is more than a "procedure for compensator design" but is rather an important part of the disturbance-feedback philosophy.

The author is in general agreement with the discussers' remarks on the effects of gain changes. As they point out, the two systems are equally sensitive to gain changes in the regulator loop compensator as far as disturbances are concerned. This in itself does not seem a serious limitation, for the plant-model positive feedback element may be made to have stable characteristics by many known methods. That the servo-response of the conditional system is independent of the regulator loop compensator is recognized, but it does not seem to be independent of changes in either the "servo model" or the "servo compensator."

The ability in the conditional design to determine a servo transient response which is different from the regulator transient response is a significant difference between the two methods. However, it is felt by the author that the response of industrial process to changes in set-point is relatively unimportant when compared with response to disturbances. The author wonders, therefore, if the ability to "tailor" the servo response independently of the disturbance response is, in general, sufficiently worth while in process control to warrant the added compensating units, especially when the individual responses produced by disturbance-response feedback, although alike, may both be near optimum.

Evaluation of the Nonlinear Kinetic Behavior of a Nuclear Power Reactor[1]

By R. C. HOWARD,[2] CAMBRIDGE, MASS.

The transient response of a nuclear-reactor system can be extremely difficult to determine by analytical methods. Frequently, recourse must be made to a computational device if a complete description of the system is to be obtained. In this study, an analog computer was employed to investigate the transient behavior of a sodium-cooled, graphite-moderated heterogeneous reactor and its primary-loop heat exchanger. The twofold purpose of the study was to analyze this particular reactor system and to obtain a general evaluation of the errors that result from the simplifying assumptions often made to solve reactor transient problems. The most important result of the study is the demonstration of the inherent short-term stability of this type of reactor and the long-term transient instability that results unless external control is employed.

Nomenclature

The following nomenclature is used in the paper:

- A' = constant relating neutron flux to power, watt sec cm^2 $\sqrt{\text{deg K}}$/neutron
- a = distance between coolant boundary and point at which average fuel temperature occurs, cm
- B^2 = geometric buckling, related to critical size of reactor, 1/cm^2
- b = distance between coolant boundary and point at which average moderator temperature occurs, cm
- C_c = total heat capacity of cold stream of heat exchanger, watt sec/deg C
- C_h = total heat capacity of hot stream of heat exchanger, watt sec/deg C
- C_i = precursor concentration; number of precursor atoms of ith type per cubic centimeter, 1/cm^3
- c_f = specific heat of fuel rod, watt sec/gm deg C
- c_h = specific heat of reactor coolant, watt sec/gm deg C
- c_m = specific heat of moderator, watt sec/gm deg C
- d_f = density of fuel rod, gm/cm^3
- d_m = density of moderator, gm/cm^3
- f = thermal utilization; ratio of thermal neutrons absorbed in fuel to total number absorbed
- G_c = mass rate of flow times specific heat for cold stream of heat exchanger, watt/deg C
- G_h = mass rate of flow times specific heat for hot stream of heat exchanger, watt/deg C
- H = height of reactor cylinder, cm
- h_f = total heat-transfer coefficient between fuel rod and coolant stream, watt/cm^2 deg C
- h_m = total heat-transfer coefficient between moderator and coolant stream, watt/cm^2 deg C
- k_{eff} = effective multiplication constant; ratio of number of neutrons produced to number lost and consumed
- k_f = thermal conductivity of fuel rod, watt/cm deg C
- k_m = thermal conductivity of moderator, watt/cm deg C
- L = thermal diffusion length; mean distance that a thermal neutron travels before destruction, cm
- L_h = total length of heat-exchanger tube, cm
- l = thermal neutron lifetime in finite medium, sec
- N = number of fuel rods in reactor
- P = power; total power generated in reactor resulting from fission energy release, watts
- p = resonance-escape probability; probability that a neutron will not be absorbed between time it is produced and time it reaches thermal energy
- R = radius of reactor cylinder, cm
- R_2 = radius of outside of coolant annulus, cm
- R_3 = radius of effective lattice cell, cm
- R_0 = radius of fuel rod, cm
- T_e = average outlet coolant temperature in reactor, deg K
- T_{ci} = inlet temperature of cold stream to heat exchanger, deg K
- T_{co} = outlet temperature of cold stream from heat exchanger, deg K
- T_f = average temperature of fuel rod, deg K
- T_{fb} = average fuel temperature at fuel-rod boundary, deg K
- T_{hi} = inlet temperature of hot sodium to heat exchanger, deg K
- T_{ho} = outlet temperature of cooled sodium from heat exchanger, deg K
- T_m = average temperature of moderator, deg K
- T_{mb} = average moderator temperature at the coolant boundary, deg K
- T_{mo} = moderator reference temperature, equal to 700 K
- T_{ri} = inlet coolant temperature to reactor, deg K
- T_{ro} = average outlet coolant temperature from reactor, deg K
- t = time, sec
- t_i = time constant of ith group of delayed-neutron precursors
- UA = product of over-all heat-transfer coefficient and its appropriate area for heat exchanger, watt/deg C
- W_h = mass rate of flow of sodium in one reactor cell, gm/sec
- β = total fraction of neutrons delayed in each generation by all precursors
- β_i = fraction of neutrons delayed in each generation by ith precursors
- ϵ = fast fission factor; ratio of number of neutrons produced by fission due to fast capture to total number of neutrons produced
- λ_i = decay constant; a measure of rate of decay of ith precursors, 1/sec
- λ_T' = decay constant for one-group delayed-neutron precursor approximation, 1/sec
- ν = average number of neutrons produced per fission-causing neutron absorbed in U-235

[1] This study is part of a thesis project supported by the U. S. Air Force, under Category II of Contract No. AF 33(616)-2263, and conducted at the D. A. C. L., Massachusetts Institute of Technology, Cambridge, Mass.

[2] Dynamic Analysis and Control Laboratory, Massachusetts Institute of Technology.

Contributed by the Instruments and Regulators Division and presented at the Instruments and Regulators Division Conference, Ann Arbor, Mich., April 25–26, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 1, 1955. Paper No. 55—IRD-8.

ρ = total reactivity $(k_{eff} - 1)/k_{eff}$
ρ_e = externally controlled reactivity
ρ_m = reactivity change due to temperature variations in reactor core
Σ_f = thermal neutron fission cross section (when multiplied by ϕ, gives number of fissions per cm³ per sec), 1/cm
Σ_{fo} = thermal neutron fission cross section at $T_m = T_{mo}$, 1/cm
τ = Fermi age; measure of distance traveled by a neutron from time of birth to the time it becomes thermal, cm²
τ_c = time taken by cold stream to flow through heat exchanger, sec
τ_h = time taken by sodium to flow through heat exchanger, sec
τ_r = time taken by coolant to flow through reactor, sec
τ_t = time taken by coolant to flow between reactor and heat exchanger, sec
ϕ = flux (not a vector quantity), neutrons/cm² sec

Introduction

The nonlinear kinetic behavior of nuclear reactors, similar to the nonlinear behavior of many complex mechanical and electrical systems, cannot be analyzed quantitatively for the general case in which all nonlinearities are included. In the unclassified literature, analyses have been presented for particular cases[3,4] in which one or more of the following simplifying assumptions has been made:

1 Prompt emission of all neutrons.
2 One group of delayed-neutron precursors.
3 No temperature dependence of reactivity.
4 Linear dependence of reactivity on temperature.
5 Power level proportional only to the flux.
6 Only small changes in power level.
7 A single, average, reactor temperature.

The results of these analyses are valid only within the limits of the assumptions upon which they are based, and the accuracy of the assumptions depends on the reactor considered and its application. The assumptions are much better suited to low, steady-power research reactors than to variable-demand power-station reactors.

This paper describes a transient analysis of a power reactor (Table 1) in which none of the foregoing assumptions was made.

TABLE 1 REACTOR DIMENSIONS AT 700 K

Height of core, H = 26.2 ft
Radius of core, R = 14.2 ft
Radius of fuel rod, R_3 = 1.5 cm
Outer radius of coolant annulus, R_2 = 1.89 cm
Radius of equivalent cell, R_1 = 10.78 cm

One purpose of the study was to investigate the error in predicting the reactor kinetic behavior that results from making these assumptions. In the course of the analysis, other discoveries were made which are of general interest. Notable among these is the unusual dependence of reactivity on temperature for the particular type of reactor chosen. This temperature dependence further complicates the nonlinear transient response of the reactor.

A realistic analysis of reactor kinetics involves two basic problems—nuclear kinetics and heat transfer. The heat transfer must be included because of its effect on the nuclear kinetics. A rise in temperature of the reactor materials alters their nuclear properties in such a manner as either to increase or decrease the rate of power rise. If the rate of power rise is decreased, the reactor is said to have a "negative temperature coefficient," and if it is increased, a "positive temperature coefficient."

Because reactors with positive temperature coefficients are unstable if not externally controlled, most reactor designs have a negative temperature coefficient. The magnitude and sign of the temperature coefficient depend on both the reactor materials and their geometrical distribution. The phrase "temperature coefficient" implies a linear variation of reactivity with temperature. The assumption of linear variation is valid for certain types of reactors or for a very small region of temperature variation for any reactor. In general, however, the actual variation in reactivity for all operating temperatures must be considered. Since this function might be highly nonlinear, an analytical solution of the reactor kinetic equations would be extremely arduous. When the variation in temperature from one region to another in the reactor is included, the reactivity is a function of several temperatures, and the hand solution of the resulting equations becomes impractical. In a problem of this nature, an analog computer is a valuable tool. With such a computer, these nonlinearities and several others can be handled without serious difficulty.

For this study, a sodium-cooled, graphite-moderated, heterogeneous reactor and the counterflow heat exchanger of the primary coolant loop, Fig. 1, were considered. The only external means of influencing reactor performance is to vary the inlet coolant temperature to the heat exchanger (T_{ci}) or the reactivity

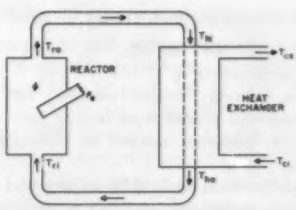


Fig. 1 Type of Nuclear-Reactor System Studied

governed by the control rod (ρ_e). These two parameters may be considered inputs, and all other variables in the problem depend on them.

Equations of the System

Several simplifying assumptions must be made if a set of usable differential equations that adequately describes the system is to be obtained. Because time is the only independent variable considered, all space distributions must be assumed constant. Therefore only the average values of space-dependent variables are considered functions of time. Other simplifications include the assumptions of a bare reactor and a constant-temperature difference between the hot and cold streams of the heat exchanger.

The equations listed here were derived from fundamental relationships with consideration given to their adaptability to solution on an analog computer. A moderator temperature-correction term for the fission cross section, which ordinarily is neglected,[5] is included in the nuclear kinetic equations for this analysis. This same correction term is used in relating neutron flux to power. The heat-transfer equations were derived from energy balances and from continuity considerations.

The nuclear kinetics were represented by the following three equations

[3] "Automatic Control of Power Reactors," by M. A. Schultz, Westinghouse Atomic Power Division, U. S. Atomic Energy Commission, AECD-3163, November 6, 1950.

[4] "Dependence of Reactor Kinetics on Temperature," by J. Chernik, Brookhaven National Laboratory, BNL 173(T-30), December 20, 1951.

[5] "The Elements of Nuclear Reactor Theory," by S. Glasstone and M. C. Edlund, D. Van Nostrand Company, Inc., New York, N. Y., 1952, chap X.

$$(\rho - \beta)\phi + \frac{\sqrt{T_m}}{\epsilon\nu\Sigma_{fo}\sqrt{T_{mo}}} \sum_{i=1}^{i=5} \lambda_i C_i = l(1-\rho)\frac{d\phi}{dt} \quad \ldots\ldots [1]$$

$$\frac{dC_i}{dt} = -\lambda_i C_i + \beta_i \epsilon \frac{\nu\Sigma_{fo}\sqrt{T_{mo}}}{\sqrt{T_m}}\phi \quad \ldots\ldots [2]$$

$$\rho = \rho_e + \rho_m \quad \ldots\ldots [3]$$

where ρ is the reactivity, β the total fraction of the fission neutrons delayed, ϕ the neutron flux, T_m the volume-average moderator temperature, ν the number of neutrons produced per fission, Σ_{fo} the macroscopic fission cross section at $T_m = T_{mo}$, T_{mo} the reference temperature equal to 700 K, λ_i the decay constant of the ith precursor, C_i the concentration of the ith precursor, l the thermal neutron lifetime, β_i the fraction of the fission neutrons delayed in the ith precursor, ϵ the fast fission factor, ρ_e the externally controlled reactivity, and ρ_m the reactivity change caused by temperature variations in the reactor core.

Equation [1] was derived from a neutron balance, Equation [2] from a precursor balance, and Equation [3] from consideration of reactivity effects. Equation [2] was repeated for each of the five groups of delayed-neutron precursors. In Equation [3], ρ_m is a complex nonlinear function of the moderator, fuel, and coolant temperatures. In computing ρ_m, many temperature effects were taken into consideration. They include over-all increase in reactor size, change in the fraction of volume occupied by fuel, coolant, and moderator, change in microscopic absorption cross sections of thermal neutrons, and density changes in fuel, coolant, and moderator. These factors influence the thermal utilization f, the resonance-escape probability p, the geometric buckling B^2, the Fermi age τ, and the thermal neutron diffusion length L^2. Because functions of more than two variables could not be instrumented on the computer, the coolant temperature was assumed to remain 15 deg C below that of the moderator in the generating of ρ_m. The computation of ρ_m, although a very time-consuming task, proved worth while; it revealed increasing reactivity with rising moderator temperature and decreasing reactivity with rising fuel temperature, Fig. 2. This type of temperature dependence complicates the transient response of the reactor.

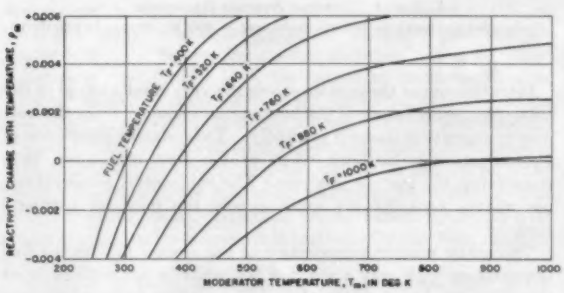


FIG. 2 TEMPERATURE DEPENDENCE OF REACTIVITY WITH T_c RELATED TO T_m BY A CONSTANT

In deriving the equations for the reactor heat kinetics, 90 per cent of the heat was assumed liberated in the fuel rods and 10 per cent in the moderator. Volume-averaged temperatures for the fuel, moderator, and coolant were considered with boundary temperatures for the fuel and moderator at their coolant boundaries. In Fig. 3 distances a and b are between the coolant boundaries and the points at which the average temperatures occur in the fuel and moderator. The heat-transfer equations employed to describe the transient heat flow within the reactor were derived to be

$$\frac{0.1P}{HN} = h_m(2\pi R_2)(T_{mb} - T_c) + d_m c_m \pi(R_3^2 - R_2^2)\frac{dT_m}{dt} \quad \ldots [4]$$

$$\frac{0.9P}{HN} = h_f(2\pi R_0)(T_{fb} - T_c) + d_f c_f \pi R_0^2 \frac{dT_f}{dt} \quad \ldots [5]$$

$$h_m 2\pi R_2(T_{mb} - T_c) = k_m 2\pi R_2\left(\frac{T_m - T_{mb}}{b}\right) \quad \ldots [6]$$

$$h_f 2\pi R_0(T_{fb} - T_c) = k_f 2\pi R_0\left(\frac{T_f - T_{fb}}{a}\right) \quad \ldots [7]$$

$$W_h c_h T_{ro}|_{t=t_0} = W_h c_h T_{ri}|_{t=t_0-\tau_r} + H h_f 2\pi R_0(T_{fb} - T_c)|_{t=t_0-(\tau_r/2)} + H h_m 2\pi R_2(T_{mb} - T_c)|_{t=t_0-(\tau_r/2)} \quad \ldots [8]$$

$$T_c = \frac{T_{ri} + T_{ro}}{2} \quad \ldots [9]$$

where P is the total reactor power; H is the fuel-rod length; N is the number of fuel rods; h_f and h_m are the over-all heat-transfer coefficients between the coolant and fuel and between the coolant and moderator, respectively; R_0, R_2, and R_3 are the dimensions shown in Fig. 3; T_{fb} is the average fuel temperature at R_0; T_{mb} is

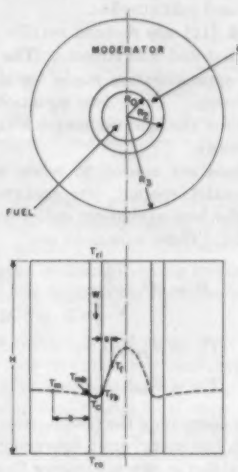


FIG. 3 AVERAGE REACTOR CELL

the average moderator temperature at R_2; T_f is the volume-averaged fuel temperature; T_c is the volume-averaged coolant temperature; T_{ri} and T_{ro} are the reactor inlet and outlet coolant temperatures, respectively; k_f and k_m are the thermal conductivities of the fuel and modulator, respectively; W_h is the coolant mass rate of flow for one cell; c_h, c_m, and c_f are the specific heats for the coolant, moderator, and fuel, respectively; d_f and d_m are the densities of the fuel and moderator; τ_r is the transport time of the coolant through the reactor; and t_0 is a dummy time variable.

Equations [4] and [5] are heat balances for the moderator and fuel, respectively. Equations [6] and [7] are heat-flow boundary conditions between the moderator and coolant and between the fuel and coolant. Equation [8] gives the enthalpy balance for the coolant, assuming that all the heat is added at the mid-plane of the reactor cylinder. Equation [9] assumes a linear temperature variation in the coolant along the cell length. Because the moderator total heat capacity was much greater than that of the fuel, and because only a small fraction of the heat was generated

in the moderator, the moderator-temperature variation was smaller and much slower than the fuel-temperature variation.

In deriving the thermal equations representing the heat exchanger, the heat capacities of the tubes were neglected, the over-all heat-transfer coefficient was assumed constant, and conduction along the stream was neglected. With these assumptions the following equations result

$$\frac{C_h}{L_h}\left(\frac{dT_{ho}}{dt}\right) + \frac{G_h}{L_h}\left(T_{ho} - T_{hi}|_{t=t_0-\tau_h}\right) + UA(T_{ho} - T_{ci}) = 0 \quad \ldots\ldots [10]$$

$$\frac{C_c}{L_h}\left(\frac{dT_{co}}{dt}\right) - \frac{G_c}{L_h}\left(T_{ci} - T_{co}|_{t=t_0-\tau_c}\right) - UA(T_{hi} - T_{co}) = 0 \quad \ldots\ldots [11]$$

where C_h/L_h and C_c/L_h are the heat capacities per unit length of the hot and cold streams; L_h is the total length of the heat-exchanger tubes; G_h and G_c are the products of the mass rates of flow and the specific heats for the hot and cold streams; UA is the over-all heat-transfer coefficient times its appropriate area; T_{ho} and T_{co} are the outlet temperatures of the hot and cold streams; T_{hi} and T_{ci} are the inlet temperatures of the hot and cold streams; τ_h and τ_c are the transport times through the heat exchanger for the hot and cold streams.

Equations [10] and [11] are derived readily from enthalpy balances made on the hot and cold streams. The second term in each equation is the approximation made for the temperature gradient along the stream. These two equations are accurate thermal representations of the heat exchanger within the assumptions mentioned previously.

Three other equations are needed to relate neutron flux to power, the reactor outlet coolant temperature to the heat-exchanger inlet, and the heat-exchanger outlet coolant temperature to the reactor inlet. These equations are

$$P = A'\frac{\phi}{\sqrt{T_m}} \quad \ldots\ldots [12]$$

$$T_{hi} = T_{ro}|_{t=t_0-\tau_l} \quad \ldots\ldots [13]$$

$$T_{ri} = T_{ho}|_{t=t_0-\tau_l} \quad \ldots\ldots [14]$$

where A' is a constant containing the fission cross section at some reference temperature, the energy-unit conversion factors, and a ratio of average fuel-rod flux to average reactor flux; τ_l is the transport time of the coolant between the reactor and the heat exchanger. The nonlinearity $1/\sqrt{T_m}$ in Equation [12] results from the $1/v$ dependence of the fission cross section where v is the velocity of the thermal neutrons.

The dependence of Σ_f on $1/\sqrt{T_m}$, which appears in Equations [1], [2], and [12], has been neglected in the past as a second-order effect. In some transients, however, the error in neglecting it is appreciable.

The equations described here represent the system model that was studied. An analog computer is well adapted to the solution of equations for this type of system.

Although approximations had to be made, these equations certainly represent a qualitative picture and, to a useful degree of accuracy, a quantitative picture of the actual system. The equations are realistic enough to determine the important parameters of the system.

Results and Analysis

Over one hundred and fifty solutions were obtained, covering about six different types of parameter variations and step changes. The effects of the rise in reactivity with increasing moderator temperature and the fall in reactivity with increasing fuel temperature result in a complex transient response. The transient is complicated further by the slow response of the moderator temperature and the relatively fast response of the fuel temperature. Each of the six situations is treated here separately.

Step Changes in Reactivity. Step changes in reactivity ranging from -0.001 to $+0.007$ were made at power levels of $\phi_0 = 1 \times 10^{13}$ to $\phi_0 = 8 \times 10^{13}$ for several cold-inlet temperatures of the heat exchanger. Typical system responses are shown in Fig. 4; the curves are replotted from oscillograms obtained during the solutions and are intended to give only a qualitative picture.

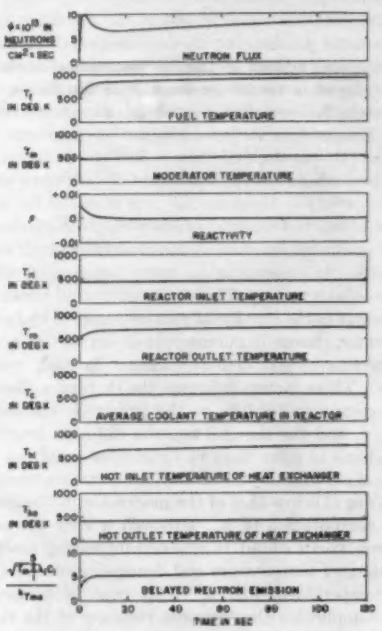


Fig. 4 Typical System Response
(Cold-inlet temperature of heat exchanger = 400 K; change in reactivity = 0.005.)

Investigation of these plots shows the very small change in the hot-outlet temperature of the heat exchanger even when the hot-inlet temperature changes markedly. This relative invariance of T_{ho} reduces the feedback effect of the heat exchanger, thus simplifying the system analysis. The cold-outlet temperature, not shown here, follows quite closely the hot-inlet temperature.

The outlet coolant temperature of the reactor follows the fuel temperature with only a slight delay which is not noticeable on the compressed scale of the plots. The average coolant temperature of the reactor follows the outlet temperature closely, and its variation is about half as large.

The fuel temperature rises rapidly, lagging slightly behind the flux, while the moderator temperature rises much more slowly than the flux. This difference in response times makes the reactivity decrease immediately after a step and then tend to increase after about 40 sec. This increase cannot be seen on the plots owing to the small scale.

Fig. 5 shows a plot of a flux response on a larger scale than that shown in Fig. 4. The nature of this response can be determined from the plot and the neutron-balance equation, Equation [1]. If $(1 - \rho)$ is assumed equal to unity, Equation [1] becomes

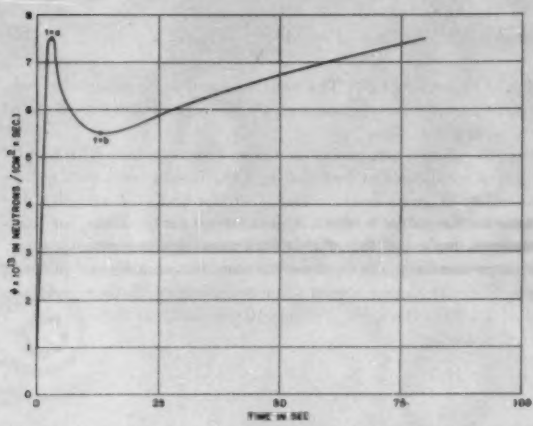


FIG. 5 FLUX RESPONSE TO LARGE STEP CHANGE IN REACTIVITY
(Cold-inlet temperature of heat exchanger = 375 K; change in reactivity = 0.006.)

$$l\frac{d\phi}{dt} = (\rho - \beta)\phi + \frac{\sqrt{T_m}}{\omega\Sigma_{fo}\sqrt{T_{mo}}}\sum_{i=1}^{i=5}\lambda_i C_i \ldots\ldots [15]$$

Before the step in reactivity, the first term on the right side of Equation [15] is equal and opposite in sign to the positive second term. At $t = 0$, a step increase in ρ is made, thus causing the negative first term to become smaller. Since the second term is unchanged, $d\phi/dt$ becomes positive, and the flux increases rapidly. Between $t = 0$ and $t = a$, ρ decreases because T_f increases. During the same time, the second term increases but lags behind the increasing flux. The decreasing reactivity and the increasing flux make the negative first term increase faster than the lagging second term. As the first term exceeds the second, $d\phi/dt$ is made zero and then negative. Therefore the flux ceases to rise, levels off at $t = a$, and then decreases. Because of the lagging second term, the reactivity does not have to be negative to cause the flux to decrease. This effect can be explained physically. Although more neutrons are produced than are consumed during each cycle and only 0.0073 of them are delayed, enough cycles occur in a very short period of time to cause a significant number of neutrons to be unavailable. Thus fewer neutrons are available to cause fission, and the flux decreases.

Between $t = a$ and $t = b$, the reactivity decreases to zero, causing the flux to decrease further. However, near $t = b$, the neutrons that were delayed during the rapid flux rise now enter the system, causing the second term to increase. At $t = b$, the second term equals the first term, and the flux levels off. If no other effects were present, the flux would settle down to a steady value after a few oscillations. However, the increase in reactivity with increasing moderator temperature prevents the flux from reaching a constant value.

After $t = b$, the effect of the large power release early in the transient causes the moderator temperature to increase appreciably, and hence the reactivity tends to rise. This increase in reactivity causes the flux to rise, which in turn causes a rise in T_f. The increase in T_f tends to cancel the increase in ρ owing to a rise in moderator temperature. The net effect of these competing factors is a tendency for the flux to level off at a high value. However, before this steady value of flux is reached, the fuel temperature exceeds a tolerable value, and the reactor may destroy itself. Therefore the reactor is not self-stabilizing for long transients. Even though there is ample time after a perturbation to compensate for the reactivity rise, the loose coupling of the system will complicate the design of an efficient automatic-control system.

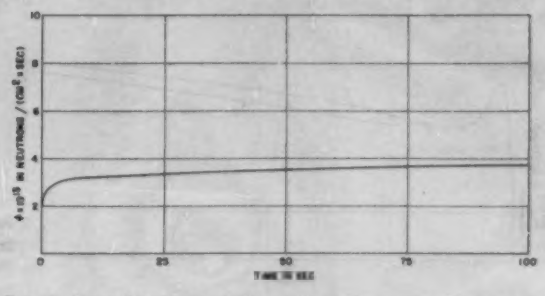


FIG. 6 FLUX RESPONSE TO SMALL STEP CHANGE IN REACTIVITY
(Cold-inlet temperature of heat exchanger = 375 K; change in reactivity = 0.0015.)

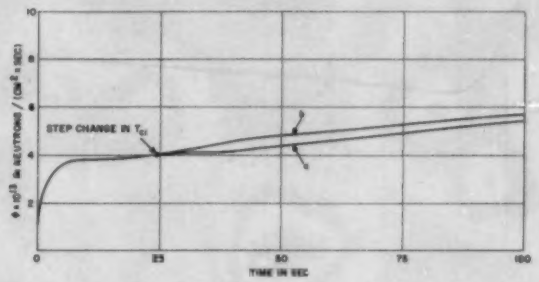


FIG. 7 FLUX RESPONSES TO STEP CHANGES IN T_{ci}
(Cold-inlet temperature of heat exchanger = 375 K; change in reactivity = 0.004.)

The response for smaller steps in reactivity is similar to that just described except for a reduction or an elimination of the initial flux peak which occurs a few seconds after the step. The long-term response again is unstable owing to the lag in the moderator temperature rise. A typical flux response to a small step in reactivity is shown in Fig. 6.

The responses studied in this phase indicate the importance of considering the temperature difference between the moderator and the fuel. If a uniform reactor temperature had been assumed, the results would have been very different.

The assumption of $T_o = T_m - 15$ C mentioned previously causes some error in the transient responses obtained. Since T_o rises rapidly above T_m by 50 to 100 C early in the response and reactivity increases with increasing T_o, the actual value of the reactivity would not decrease as quickly as observed in this study. Thus the initial flux peak would be higher. It is doubtful that the inclusion of this effect would change the qualitative picture of the reactor's dynamic performance.

Step Changes in T_{ci}. During the transient following a step change in reactivity, a step change in the cold-inlet temperature to the heat exchanger was simulated. When this temperature was decreased suddenly, T_{ho}, T_{ri}, and T_f decreased in that order. The decrease in T_f caused a rise in the reactivity and hence in the flux. An increase in T_{ci} had exactly the opposite effect. Both cases are shown in Fig. 7 for responses having identical initial conditions. Curve *a* represents an increase in T_{ci} of 40 C, and curve *b*, a decrease in T_{ci} of 27 C. The flux change occurs about 4 to 5 sec after the change in T_{ci} owing to the transport delays and the temperature lags.

Step Changes in T_{ri}. During the transient following a step change in reactivity, a step change in the inlet coolant temperature to the reactor was simulated. The effect of this change was similar to that of stepping T_{ci}. When T_{ri} was suddenly decreased, the fuel temperature fell, causing an increase in the reactivity and

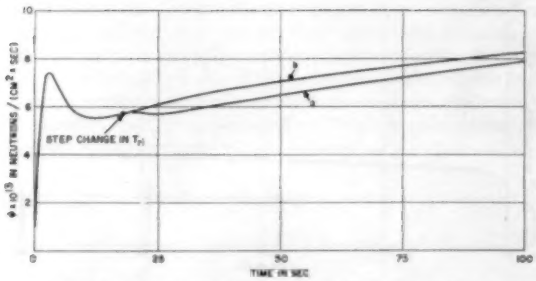


FIG. 8 FLUX RESPONSES TO STEP CHANGES IN T_{ri}
(Cold-inlet temperature of heat exchanger = 375 K; change in reactivity = 0.006.)

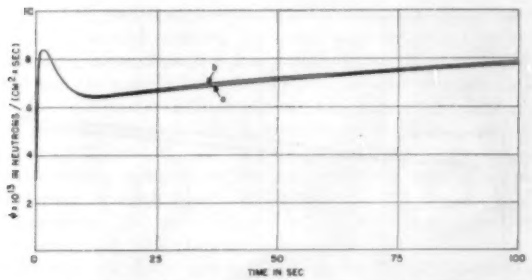


FIG. 9 FLUX RESPONSES SHOWING EFFECT OF MAINTAINING T_{ri} CONSTANT
(Cold-inlet temperature of heat-exchanger = 375 K; change in reactivity = 0.004.)

hence an increase in the flux. The opposite occurred when T_{ri} was increased. Fig. 8 shows both cases. In curve *a*, T_{ri} was increased by 56 C, and in curve *b*, T_{ri} was decreased by 28 C.

The results of an attempt to check a power rise by "cooling off" the reactor are interesting to note. With the strong, short-term, negative temperature gradient of reactivity for this type of reactor, this procedure would have disastrous results. This effect, sometimes referred to as the "cold-water accident," could occur accidentally in power reactors containing separate coolant loops.

Variation in Coolant Transport Time. The coolant transport time τ_t was varied during the study from zero to six sec. Although in a complete power system this transport time might be important, in the system studied and in the types of responses obtained, the transport time had very little effect on the flux response. The principal reason for this insensitivity to transport time was the small slow change in T_{ri} during a transient. Fig. 9 compares two runs for identical conditions except that in *b*, T_{ri} was held constant at its initial value. Holding T_{ri} constant has the same effect on the flux that an infinite transport time would have. The small difference between the responses is apparent. The differences for variations in transport time from zero to six sec are too small to distinguish and, therefore, plots of these runs are not included here.

In a complete power system where changes occur in the cold stream of the heat exchanger owing to load variations, the coolant transport time would determine how quickly the reactivity was affected and, therefore, could be a critical parameter.

Temperature Dependence of the Fission Cross Section. In most analyses the power is assumed to be related to the thermal neutron flux by a constant factor. However, included in this factor is the fission cross section which is not constant but inversely proportional to the square root of the neutron temperature. If the neutron temperature is assumed to be that of the moderator, the following relation exists

$$P = A' \frac{\phi}{\sqrt{T_m}} \quad \ldots\ldots\ldots\ldots\ldots\ldots [16]$$

which is Equation [12]. The neutron-balance equation also contains the fission cross section and, therefore, includes this moderator temperature effect.

Fig. 10 shows a comparison between two responses with identical initial conditions except that in *b* the fission cross section was held constant at its initial value. At the end of 7 min the difference between the curves is about 20 per cent. Thus, for long-term transients, neglect of this effect would cause serious error. For short-term transients, the error would be smaller and perhaps negligible. The error varies almost linearly with time. At the end of 3.5 min, the error is about 10 per cent; at the end of 1.75 min, 5 per cent.

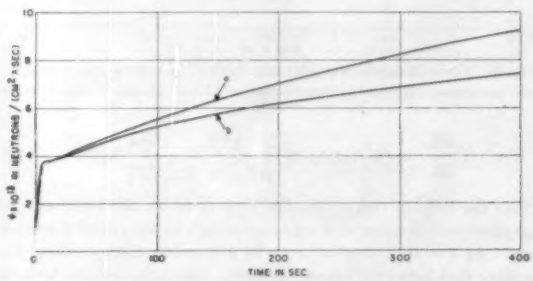


FIG. 10 FLUX RESPONSES SHOWING EFFECT ON FLUX RESPONSE OF TEMPERATURE DEPENDENCE OF FISSION CROSS SECTION
(Cold-inlet temperature of heat exchanger = 375 K; change in reactivity = 0.004.)

In a power rise, the delayed-neutron precursors are produced at a rate proportional to the value of the fission cross section. However, when the delayed neutrons enter the system, the fission cross section has decreased. This decrease makes the rate of production of prompt neutrons relatively less by the square root of the ratio of the moderator temperature at the time that the delayed-neutron precursors are produced to the temperature at the time the delayed neutrons enter the system. Thus the relative importance of the delayed neutrons in the neutron-balance equation becomes greater by this same ratio. The relative increase in the delayed-neutron production causes the flux to rise more rapidly and partially accounts for the difference between the curves in Fig. 10.

The power dependence on fission cross section tends to maintain the reactor temperatures at lower values, which causes the reactivity, and hence the flux, to be slightly higher. This power dependence accounts for the rest of the difference in the curves of Fig. 10.

For complete power-system control studies in which the transient times are long, the neglect of the temperature dependence of fission cross section would result in about 20 per cent error. For safety studies where the transient times are of the order of milliseconds, this temperature dependence is probably negligible.

Delayed-Neutron Approximations. As indicated previously, a one or two-group approximation to the five groups of delayed-neutron precursors often is made in the study of reactor transients. In order to evaluate the accuracy of these assumptions, a one-group approximation was made for several solutions. The equivalent decay constant for the one group of precursors was obtained from

$$\lambda'_T = \frac{\beta}{\sum_{i=1}^{i=5} \beta_i t_i}$$

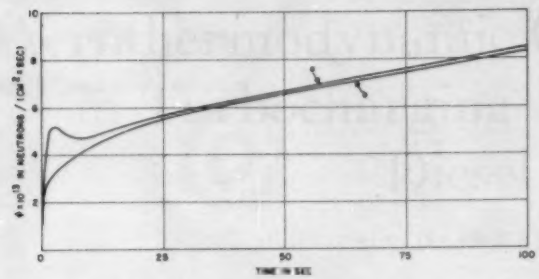


FIG. 11 FLUX RESPONSES SHOWING EFFECT OF A ONE-GROUP APPROXIMATION FOR DELAYED-NEUTRON PRECURSORS
Cold-inlet temperature of heat exchanger = 375 K; change in reactivity = 0.005.)

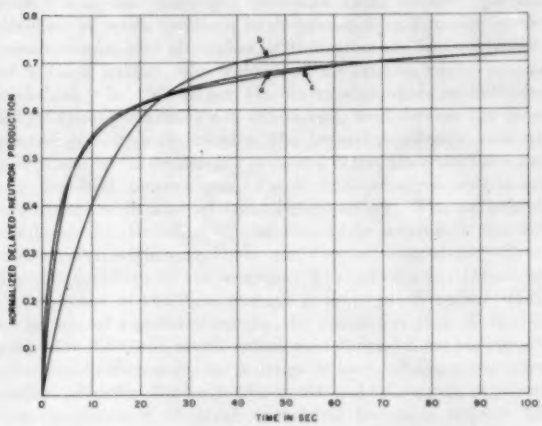


FIG. 12 CALCULATED DELAYED-NEUTRON RESPONSES TO STEP CHANGES IN FLUX TO EVALUATE APPROXIMATIONS FOR DELAYED-NEUTRON PRECURSORS

Fig. 11 shows the marked difference in the flux response resulting from this approximation. Curve *a* shows the flux response when all five groups of delayed-neutron precursors are included; curve *b* is the flux response when the one-group approximation is made. The elimination of the initial peak in curve *b* is caused by the much smaller early increase in the term for delayed-neutron production in Equation [15]. Use of the one-group decay constant makes the delay time of the neutrons that would have entered the system early in the transient unrealistically long. If no nonlinear temperature effects were included in the equations, the one-group approximation would give correct asymptotic results. However, because the temperatures are proportional to the time integral of the flux and the reactivity is proportional to the temperatures, even the asymptotic value of the flux will not be correct if the one-group approximation is used.

Accuracy is improved greatly if two delayed-neutron groups are used. One of the groups has a smaller time constant than the single delayed-neutron group, the other, a larger time constant. By this means, sufficient delayed neutrons are emitted early in the transient to cause the flux to experience the initial peak and to preserve the asymptotic accuracy. Although responses were not taken using the two delayed-neutron groups, it is easy to demonstrate the gain in accuracy resulting from this procedure.

In Fig. 12 responses of the delayed neutrons to a step in the neutron flux are shown for the following three cases: *a* the full five groups of delayed neutrons; *b* the one-group approximation; *c* the two-group approximation. The two-group approximation probably would be accurate enough for most transient analyses.

CONCLUSIONS

Several conclusions may be drawn about a sodium-graphite power reactor and the applicability of certain simplifying assumptions to investigate its transient analysis.

1 The reactivity decreases nonlinearly with fuel temperature and increases nonlinearly with moderator temperature.

2 The reactor is self-compensating for large steps in reactivity early in the transient, but exhibits an instability late in the transient if not externally controlled.

3 The entire reactor power system is very loosely coupled and could represent a difficult problem in control.

4 A step decrease in the reactor inlet temperature or in the cold stream of the heat exchanger causes the flux to rise.

5 The effect on the flux of the coolant transport time between the reactor and heat exchanger is negligible for step changes in reactivity.

6 The neglect of temperature dependence of fission cross section in the neutron-balance equation and in the flux-power relationship causes appreciable error in transients involving large temperature changes.

7 A one-group approximation to the delayed-neutron precursors causes large error early in the transient and appreciable error late in the transient. A two-group approximation should be accurate enough for most analyses.

8 For this type of reactor, a transient analysis that assumes only a single average reactor temperature would be of doubtful value.

9 The assumption of a linear dependence of reactivity on temperature would be a good approximation for small variations in power level if the appropriate linearization were performed. For any other conditions, however, this assumption of linear dependence would lead to considerable error.

ACKNOWLEDGMENT

The author wishes to express his appreciation to Dr. M. Benedict, supervisor of the thesis upon which this paper is based, for his valuable advice and comments during the solution, analysis, and write-up of the problem. The writer is indebted also to Dr. J. A. Hrones, director of the Dynamic Analysis and Control Laboratory, and Dr. W. W. Seifert, assistant director, for making the facilities of the D.A.C.L. available for this investigation. Thanks are due Mr. J. M. Gallagher, Jr., for his assistance during the simulator program and helpful suggestions in the analysis and Mr. J. Eige for his help in obtaining a description of the transient thermal behavior of the system.

Aerothermodynamic Considerations Involved in Turbocharging Four and Two-Cycle Diesel Engines

By RUDOLPH BIRMANN,[1] TRENTON, N. J.

The turbocharged engine is, in effect, a compound engine in which the engine itself handles the high-pressure end of both compression and expansion, while the exhaust-turbine-driven compressor takes care of the low-pressure end. The turbocharger pressure ratio determines the division of work between high-pressure and low-pressure components, and the value of this pressure ratio governs, to a large extent, the values of all engine performance variables. In this paper the interrelationships between all of these variables are discussed, and curves are presented giving, numerically, the interdependence, one on the other, of turbocharger pressure ratio, thermal loading, specific fuel consumption, and turbocharger combined efficiency, with and without intercooling. The aerothermodynamic situation of the two-cycle engine is treated separately, with emphasis on the all-important flow-handling ability of the engine. This ability is defined by the concept of the "gas change process merit ratio," and, by means of a series of curves, the specific engine air flow is given for different merit ratios and different turbocharger combined efficiencies for a range of turbocharger pressure ratios. Finally, for both two-cycle and four-cycle engines, the importance of good matching between engine air consumption characteristics and turbocharger air delivery characteristics is discussed, and as a criterion for this matching the engine operating line, as it appears superimposed on the compressor characteristic curves, is used.

Introduction

THE beginning of turbocharging the reciprocating internal combustion engine dates back to before World War I. It was not until the present decade, however, that the turbocharger became an essential component, and even then, only for large and medium-size four-cycle diesel engines. This recent conversion to turbocharging was inevitable, once it had been proved that the large increase in output and the substantial improvement in fuel consumption brought about by turbocharging makes the atmospheric-aspirated engine obsolete by leaving no economic justification for its existence.

Despite this tremendous progress in turbocharging large and medium-size diesel engines—progress that has resulted in the achievement of bmep as high as 300 psi, and fuel consumptions as low as 0.3 lb/bhp hr—turbocharging of the small automotive-type engine is virtually nonexistent, and what little there is does not take advantage of the advances that have been made in the turbocharging of large engines (for example, the use of high turbocharger pressure ratios—2.0 and above). This situation may be due to the fact that neither high-performance nor high-pressure turbochargers of the required small capacity have been available to the manufacturers of small engines, and also to the fact that these manufacturers do not appreciate the full extent of the advantages to be derived from turbocharging.

Because the aerothermodynamic considerations involved in the turbocharging of four-cycle and two-cycle engines are entirely different, the first section of this paper will be confined to the turbocharging of four-cycle engines, and that of the two-cycle engine will be treated separately later on. It should be pointed out, however (before embarking on a detailed discussion of the improvements that turbocharging can make in the operating characteristics and performance of both two and four-cycle engines), that the improvements to be obtained *are limited solely by the ability of the engine to handle the increased mechanical and thermal loadings involved*, and it might be well to define these terms as they will be used herein.

Mechanical Loading

Mechanical loading is measured in terms of the stresses in all engine parts—particularly those stresses that are produced by the gas pressures in the cylinders—and in terms of all the bearing loadings affected by these gas pressures. In very large low-speed engines an appreciable increase in the mechanical loading often cannot be tolerated, and turbocharging therefore must be accomplished in such a manner that the internal pressure level remains the same, or very nearly the same. This can be done by reducing the engine compression ratio, or by means of a suitable valve and fuel-injection timing, or by the use of the well-known Miller cycle, the objective of which is to make possible a high external pressure level (that is, high turbocharger pressure ratio) without a substantial increase in the internal pressure level (high compression and firing pressures).

In small high-speed engines the peak loading of the bearings is often brought about by inertia forces, rather than by pressure forces (as in large low-speed engines), and therefore increasing the internal pressure level does not necessarily increase the maximum bearing loading. Similarly, small engines can more easily handle increased internal pressures, from the point of view of cylinder wall stresses, because the thicknesses of the cylinder walls are usually chosen from considerations of casting and manufacturing requirements, instead of being dictated by the allowable stresses.

Thermal Loading

Thermal loading is measured in terms of the rate of heat flow through all the walls enclosing the engine cylinder space (that is, the rate of heat flow from the cylinder charge to the cooling mediums—which usually are water and oil—through the cylinder liners, the cylinder heads, and the piston bottoms). In large engines this rate of heat flow cannot be increased without shortening the life of the engine and endangering its reliability. In small en-

[1] Consulting Engineer, De Laval Steam Turbine Company. Mem. ASME.

Contributed by the Oil and Gas Power Division and presented at the Oil and Gas Power Conference, Washington, D. C., June 5–10, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, April 6, 1955. Paper No. 55—OGP-10.

gines, where the ratio of total heat-transfer surfaces to cylinder volume is more favorable, a limited increase in thermal loading often can be permitted.

In general, it cannot be overemphasized that thermal loading has a much more drastic effect on the mechanical performance of the engine than has mechanical loading. Experience has shown that many failures in large engines (such as ring and groove troubles, cylinder head and piston bottom cracking, and liner scoring), which at one time were ascribed to excessive firing pressures, in reality were due to a thermal loading that was too high. The output of many a large engine has been increased by turbocharging—up to as much as 250 per cent of its atmospheric aspirated rating—without any adverse effect on engine life and reliability, because this greatly increased output was obtained without an increase in the thermal loading, even though the mechanical loading was stepped up far above that originally considered allowable by the designer and manufacturer.

From basic heat-transfer considerations, the momentary heat flow from the cylinder charge (momentary temperature T) to the cooling water (temperature T_c) is

$$q = K(T - T_c) \quad [1]$$

The over-all heat-transfer coefficient K is

$$K = \frac{1}{\frac{1}{\alpha_1} + \frac{\delta}{\lambda} + \frac{1}{\alpha_2}} \quad [2]$$

where

α_1 = heat-transfer coefficient, gas to cylinder wall
λ = heat-conductivity factor of cylinder wall
α_2 = heat-transfer coefficient, cylinder wall to cooling medium
δ = thickness of cylinder wall

Since α_2 is very large, $1/\alpha_2$ can be set as zero. δ/λ is a known constant, which we shall call c, and relatively small in comparison to $1/\alpha_1$.

Professor Eichelberg has shown that α_1—that is, the coefficient of the heat transfer between gas and cylinder wall—is proportional, at any one instant, to $\sqrt{pT}$, where p is the momentary pressure, and T the momentary gas temperature.

If k is the proportionality factor, then

$$\alpha_1 = k\sqrt{pT} \quad [3]$$

Now, therefore

$$K = \frac{1}{\frac{1}{k\sqrt{pT}} + c} \quad [4]$$

and q, in view of Equation [1], becomes

$$q = \frac{1}{\frac{1}{k\sqrt{pT}} + c}(T - T_c) \quad [5]$$

For any given engine, the total heat flow Q (in Btu per minute) is known from tests. (The total heat flow Q is the sum of the instantaneous flows, or, in other words, $Q = \Sigma q$.) These tests also permit plotting the entire internal cycle of the engine in the temperature-entropy diagram, which makes it possible to determine the pressure p and the gas temperature T at any one instant, and use these to determine from Equations [1] and [4] the instantaneous heat flow to the cooling medium. By trial and error the coefficient k, in Equation [5], is then adjusted, and a value for k is found that gives the amount of the heat flow actually measured.

Basic Thermodynamic Relationships in Turbocharged Four-Cycle Engines

Analytical Approach. Having determined the value of k that applies to a definite internal cycle of a given engine—that is, for operation of this particular engine under known conditions of inlet manifold pressure and temperature, air-fuel ratio, and measured heat rejection to the cooling mediums—any number of hypothetical cycles can be investigated and the effects of varying turbocharger pressure ratio and/or air-fuel ratio on the thermal loading of the engine and its bmep can be accurately predicted, for which purpose it is necessary to plot the temperature-entropy diagram for each modified cycle—that is, for each new set of conditions of inlet manifold pressure, temperature, and air-fuel ratio. Alternately, it can be specified that the thermal loading must be constant, and for a given inlet manifold pressure and temperature the air-fuel ratio can be determined that, for the known value of k, gives the specified value of the heat flow to the cooling mediums.

Particularly since this process is largely one requiring a trial-and-error approach, it involves much tedious and time-consuming computation. Nevertheless, the results are well worth the effort, since they give a clear insight into the interrelationship of the many variables involved. For this reason these computations have been performed, and the work has been carried out in such a way that the results are indicative of the performance of the average engine-and-modern-turbocharger combination. The computations were based on an engine compression ratio of 13.5:1, a turbocharger combined efficiency of 60 per cent, a turbocharger compressor efficiency of 80 per cent, an ambient air temperature of 90 F, and, where intercooling is employed, an intercooler "out" temperature of 100 F.

Results of Analysis. The results of this study, presented in Fig. 1, are in the form of curves—all having turbocharger pressure ratio as abscissa. They show the effect of various turbocharger pressure ratios on the bmep of the engine, the air-fuel ratio, the thermal loading, and the specific fuel consumption—for intercooling to 100 F (solid-line curves), as well as without intercooling (dash-line curves). In the first vertical series of curves—(A) through (D)—the air-fuel ratio of the engine is maintained constant; in the second vertical series—(E) through (H)—the thermal loading is maintained constant; and, finally, in the remaining series of curves—(I) through (L)—the bmep is maintained constant. A fourth series of curves could be plotted to show the effect of holding the mechanical loading (or firing pressure) constant. This was not done, nor were curves of peak pressures plotted, because (as mentioned previously) the majority of engines (especially small high-speed engines) are capable of handling substantially increased mechanical loadings, particularly if this increase is not accompanied by an increase in the thermal loading, or only a moderate increase is permitted. Alternately, in any new engine development, provisions can be made to cope with the required high internal pressure level.

The three series of curves in Fig. 1 afford an interesting comparison. It will be noted, for example, that the greatest increase in bmep (with increasing turbocharger pressure ratio) can be obtained with constant air-fuel ratio; however, this increase is associated with a very substantial increase in the thermal loading. This is brought about by the fact that with constant air-fuel ratio the mean cycle temperature does not change with increasing turbocharger pressure ratio. However, the charge density increases, and the heat transfer (from the charge, through the cylinder walls, to the cooling medium) is consequently increased. In other words, the thermal loading of the engine is "stepped up," as shown by the thermal loading curves, which are based on the

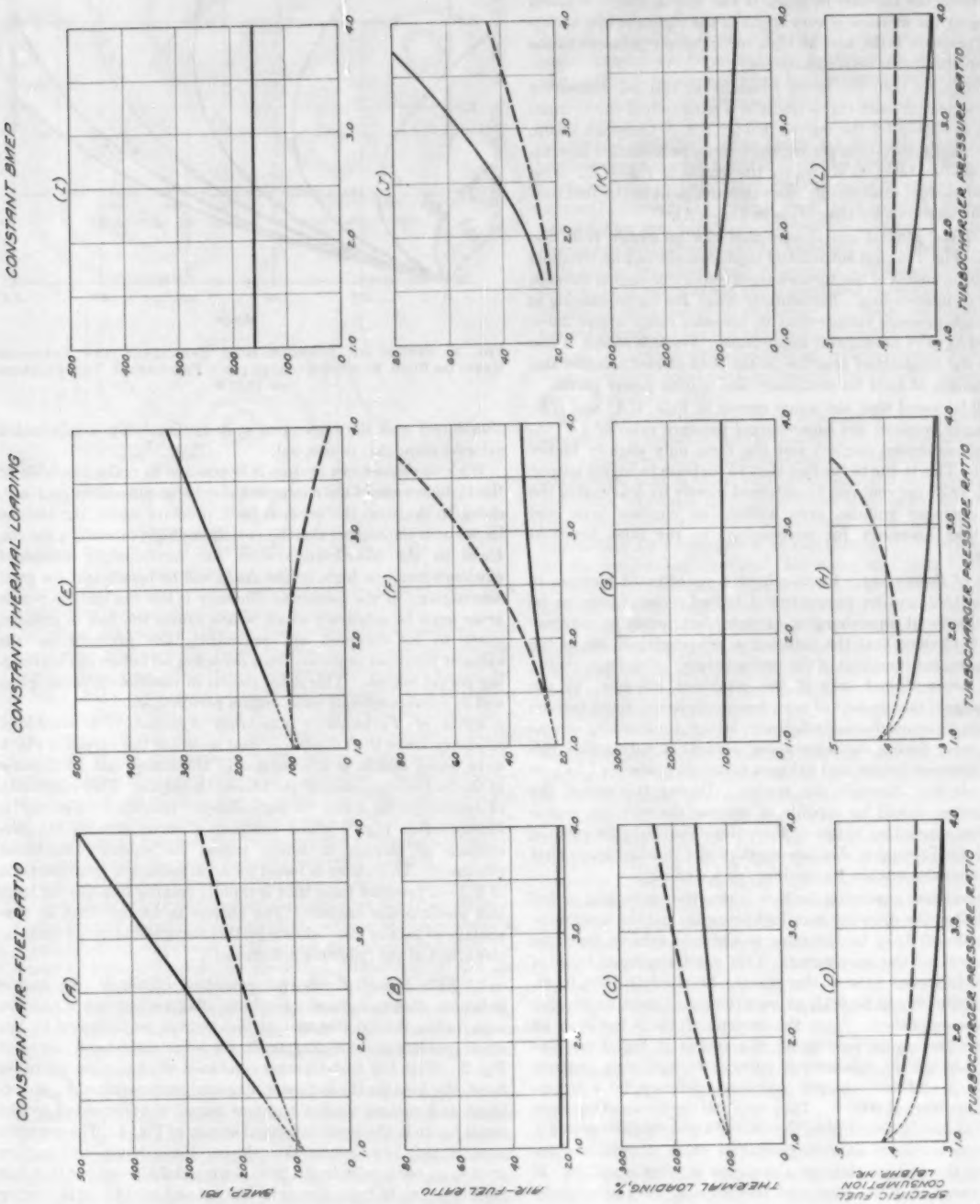


FIG. 1 THERMODYNAMIC RELATIONSHIPS IN FOUR-CYCLE ENGINES

thermal loading of the atmospheric-aspirated engine being 100 per cent.

If the thermal loading is maintained constant and intercooling is employed, the increase in bmep is less spectacular. Without intercooling the increase is very slight in the region of low turbocharger pressure ratio, and at high turbocharger pressure ratios the bmep obtainable even diminishes.

If, finally, we hold the bmep constant at 160 psi, increasing turbocharging pressure ratios result in a reduction of the thermal loading, as shown by the curves in Fig. 1(K). Constant bmep, with increasing turbocharger pressure ratio, necessitates increasing the air-fuel ratio, as shown by the curves in Fig. 1(J). This increased air-fuel ratio results in an improvement in the fuel consumption, as shown by the curves in Fig. 1(L).

One of the general conclusions that can be drawn from the curves in Fig. 1 is that substantial improvement can be obtained in the performance of the turbocharged four-cycle engine through the use of intercooling. Particularly when the turbocharging is in the high-pressure range—that is, pressure ratios above 2.0—intercooling is of paramount importance. For this reason it has become the established practice in the field of medium-size and large engines, at least for stationary and marine power plants.

It will be noted that the bmep curves in Figs. 1(A) and 1(E) rise sharply between the supercharger pressure ratio of 1.0 (atmospheric-aspirated engine) and the ratio only slightly higher than 1.0. This is due to the fact that an increase in engine output as much as 25 per cent can be obtained merely by scavenging the engine clearance volume, even without an increase (over and above that necessary for scavenging) in the inlet manifold pressure.

Effect of Scavenging. Inasmuch as a pronounced increase in engine output and an improvement in fuel consumption can be achieved without supercharging, as such—but merely by scavenging—it is obvious that the inducing of scavenging is one of the most important functions of the turbocharger. Good scavenging can be accomplished only if the combined efficiency of the turbocharger (the product of compressor efficiency, times turbine efficiency, times mechanical efficiency) is high, thus making available (at least during the scavenging period) a substantial difference between intake and exhaust manifold pressures (Δp), to induce air flow through the engine. During this period the turbocharger should be capable of passing through the engine an air flow amounting to two or three times the clearance volume, in order to thoroughly displace most of the residual gases that would otherwise remain trapped by fresh cool air.

The basic laws governing the flow during the scavenging period demand that for different manifold pressures (p) the nondimensional ratio of $\Delta p/p$ be the same in order to achieve the same effectiveness of the scavenging. This nondimensional ratio of pressure difference (across the engine) to absolute inlet manifold pressure should be 0.15, or even larger, if good scavenging is to be accomplished. Since the concept of $\Delta p/p$ has such an important bearing on scavenging, the curves in Fig. 2 are presented, showing the relationship between turbocharger pressure ratio, $\Delta p/p$, and turbocharger combined efficiency, for a preturbine temperature of 1000 F. They apply to the constant-pressure system (as distinguished from the blowdown or "pulse" system), or, in other words, to an arrangement in which all cylinders discharge to the turbine through a common exhaust manifold. It will be noted that the combined efficiency of the turbocharger must be sufficiently high, and also that the turbocharger pressure ratio must have a certain minimum value, in order to insure good scavenging. With very low turbocharger pressure ratios good scavenging cannot be obtained, even with the highest possible turbocharger efficiencies, unless the single manifold system is abandoned and the blowdown system (involving a subdivided exhaust manifold) is adopted.

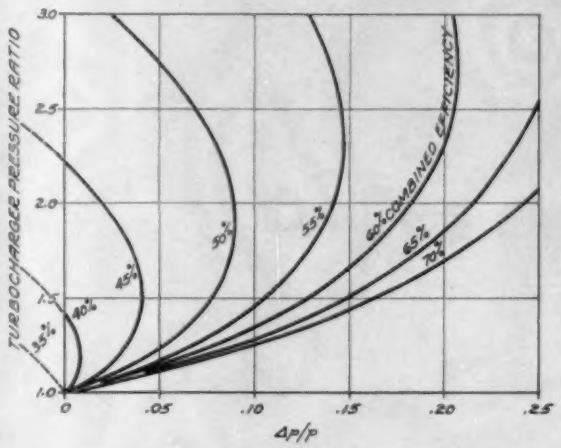


FIG. 2 EFFECT OF TURBOCHARGER EFFICIENCY AND PRESSURE RATIO ON FLOW FUNCTION ($\Delta p/p$) FOR A PRETURBINE TEMPERATURE OF 1000 F

With the blowdown system it is possible to make available to the turbine some of the energy in the exhaust pulsations, and in so doing to decrease the exhaust back pressure which the turbine imposes on the engine, thereby correspondingly increasing the Δp. Even in the blowdown system the turbocharger combined efficiency must be high, or the $\Delta p/p$ will be insufficient for good scavenging. If the combined efficiency is low the turbine nozzle areas must be relatively small, which causes the flow of exhaust gases to be dammed up, preventing the pressure in the exhaust manifold branches from dropping off before the scavenging period begins. This often results in insufficient scavenging, and in correspondingly poor engine performance.

Effect of Turbocharger Combined Efficiency. A combined efficiency lower than the 60 per cent on which the curves of Fig. 1 were based results in a reduction of the bmep and an increase in the fuel consumption of the four-cycle engine. This sensitivity of engine performance to turbocharger combined efficiency is illustrated in Fig. 3 which presents a curve showing the percentage of change in bmep versus turbocharger combined efficiency. This curve is based on a turbocharger pressure ratio of 3:1—a pressure ratio that is rapidly coming into use for large and medium-size engines. The change in bmep—that is, percentage of gain or loss—shown by this curve is due to the cumulative effect of the following influences:

1 With high turbocharger combined efficiency, the engine indicator diagram (pressure-volume diagram) shows a positive area during the exhaust and suction strokes, as indicated by the small pressure-volume diagram in the lower right-hand corner of Fig. 3. With low turbocharger combined efficiency, on the other hand, the loop (in the indicator diagram) circumscribed by the exhaust and suction strokes becomes negative, as indicated by the small figure in the lower left-hand corner of Fig. 3. The negative area of this loop means that during the exhaust and suction strokes power is withdrawn from the crankshaft—power that has been taken away from the engine output and used to aid in driving the turbine, by pneumatic transmission thereto.

2 With a negative Δp, or even with a small Δp (resulting from a low turbocharger combined efficiency), more residual exhaust gases remain in the cylinders, causing dilution of the fresh charge and thereby reducing the amount of fresh air trapped in

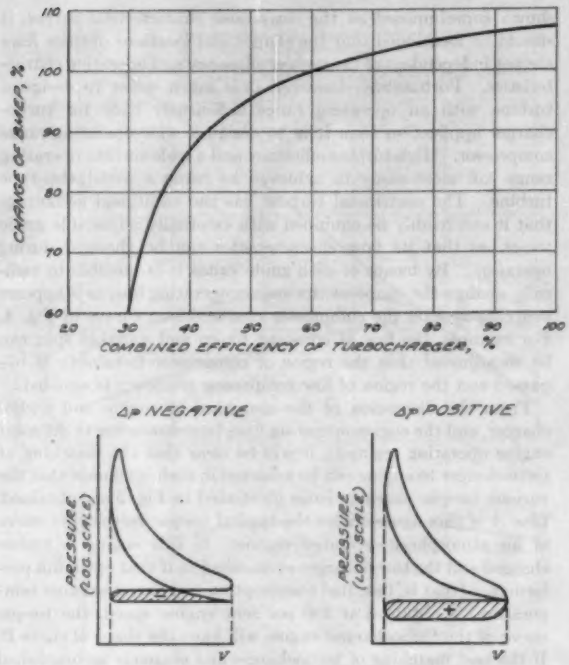


FIG. 3 EFFECT OF TURBOCHARGER COMBINED EFFICIENCY ON ENGINE BMEP FOR A TURBOCHARGER PRESSURE RATIO OF 3.0

the cylinders. Consequently less fuel can be burned, and as a result there is a reduction in engine output.

3 Poor scavenging (due to poor combined efficiency), and therefore dilution of the fresh charge by the residual exhaust gases, results in a pronounced temperature increase at the beginning of the compression stroke. (Incidentally, if intercooling is not employed, this heating can be aggravated seriously by a low compressor efficiency.) This increased temperature at the beginning of the compression stroke increases the work of compression, lowers the cycle efficiency, and reduces the net output—an effect that can easily be recognized if it is remembered that the net output of any internal combustion engine is the relatively small difference between the large amount of work required for compression and the work derived from the expansion.

Improved Engine Thermal Efficiency, Through Turbocharging. While turbocharging is generally thought of as the most effective means for increasing the output of the engine at the expense of relatively little additional weight and manufacturing cost, the curves of Figs. 1(*D*) and (*H*) show that turbocharging also improves the fuel consumption, particularly if intercooling is employed. Tests of many large engines have demonstrated that the improvement in fuel consumption alone will allow the cost of the turbocharger to be paid for out of fuel savings in a period of less than two years.

The better fuel consumption of the turbocharged engine is due to several factors. First, its mechanical efficiency is higher than that of the atmospheric-aspirated engine, simply because the mechanical losses increase very little with increased engine output, and, percentagewise, become much smaller. The second factor leading to improved fuel consumption of the turbocharged engine, particularly if intercooling is used, is that, percentagewise, less of the total heat in the fuel is rejected to the engine cooling system, leaving more of the fuel heat to be converted into net engine output. This effect can be quite drastic, reducing the heat rejection to the cooling system from as much as 20 per cent of the heat in the fuel, in the atmospheric-aspirated engine, to as little as 8 per cent in the same engine turbocharged. Finally, the turbocharged engine operates with a higher combustion air excess, which improves the thermal efficiency of the diesel cycle, as will be discussed.

Higher Air-Fuel Ratios, Through Turbocharging. Increased engine output and improved fuel consumption are not the only advantages to be derived from high-pressure turbocharging. In some cases, where an increase in bmep is not needed but a lowering of the mean temperature level of the engine and an increase in the air-fuel ratio for a given bmep are desired, going to higher turbocharger pressure ratios is again the answer. This statement is substantiated by the third vertical series of curves of Fig. 1—curves (*I*) through (*L*)—which apply to a constant bmep of 160 psi and show that with high-pressure turbocharging a tremendous increase in air-fuel ratio can be achieved. This increased air-fuel ratio lowers the mean cycle temperature to a considerable extent even without intercooling, and a great deal more when only a small degree of intercooling is employed. A large air-fuel ratio (for normal operation) has the advantage of permitting rapid acceleration of the engine without smoking, which is an important consideration in a diesel-engine-powered vehicle—for example, in a locomotive diesel engine.

An even more important effect of a high air-fuel ratio is the improvement it makes in the thermal efficiency of the internal cycle, over and above the increase in efficiency resulting from the better combustion that is often associated with the higher combustion air excess. In large and medium-size engines it has been found that poor combustion, and consequently smoking, occurs if the combustion air excess factor is less than 1.5. Between 1.5 and 2.0 the combustion constantly improves, and from 2.0 on up no further improvement in combustion can be achieved; nevertheless, there is an improvement (with high combustion air excess factors) in the thermal efficiency of the internal cycle. In this connection, it is interesting to philosophize that with an infinitely high combustion air excess factor, the efficiencies of the Carnot cycle and the diesel cycle are the same.

Advantages of Intercooling. Fig. 1 brings out clearly the advantages of intercooling. These advantages—in terms of increased bmep, reduced thermal loading, a larger air-fuel ratio, and improved fuel consumption—are so great that intercooling, as previously stated, has become the established practice in the field of medium-size and large engines, at least for stationary and marine power plants. Even for mobile installations it is inevitable that intercooling will become the rule, rather than the exception. Serious consideration is even being given to the use of air-to-air intercoolers for vehicle power plants, which seems logical, since in these installations all cooling heat must ultimately be dissipated to the ambient air, and, even more important, for a given bmep output the amount of heat that is removed from the system by intercooling does not have to be removed by the jacket-cooling-water radiator. This interesting observation, which is not always clearly understood, means that in a vehicle powered by a turbocharged diesel engine the heat load of the coolant radiator is reduced by an amount nearly as large as the heat-dissipation capacity of the air-to-air intercooler.

OPERATING CHARACTERISTICS OF TURBOCHARGED FOUR-CYCLE ENGINES

Fig. 4 presents, in nondimensional fashion, the characteristic curves of a modern, highly efficient turbocharger compressor. The basic curves (solid lines) relate inlet volume flow (in per cent) to turbocharger pressure ratio. Superimposed on these curves are contour lines of constant adiabatic compressor efficiency. It will be noted that the turbocharger compressor is capable of oper-

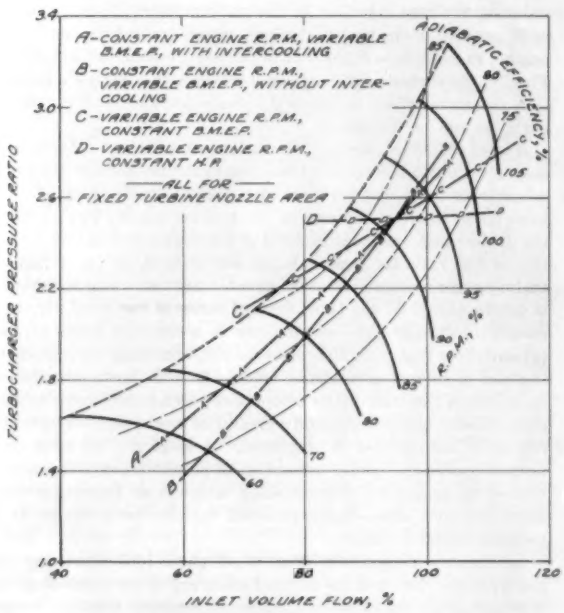


FIG. 4 TURBOCHARGER COMPRESSOR CHARACTERISTICS AND FOUR-CYCLE ENGINE OPERATING LINES

ating, under conditions of maximum efficiency, over only a relatively narrow range. For this reason it is very important that the turbocharger flow characteristics be matched accurately to the air consumption characteristics of the engine. In other words, even if the turbocharger efficiency is the highest possible, the engine performance will not be good unless the turbocharger is properly matched to the engine. It has been found that the best way to judge whether or not there is good matching is to superimpose the engine operating line on the compressor characteristic curves.

The engine operating line of a well-matched turbocharger-engine combination, in terms of turbocharger pressure ratio and per cent of free air volume flow, is shown by line *A* in Fig. 4. This operating line corresponds to constant engine rpm, intercooling, and variable engine load (or bmep). Without intercooling, the slant of the engine operating line is more steep—approximately as shown by line *B*. Line *C* shows the approximate engine operating line for conditions of constant bmep and variable engine rpm, whereas line *D* shows an operating line for constant engine horsepower under conditions of variable rpm. It will be noted that in so far as the compressor goes, good matching—that is, high compressor efficiency—is obtained over the entire extent of the constant-speed, variable bmep operating range, whereas with operation at constant bmep the operating line traverses the compressor high efficiency region, and at high engine loads—that is, high turbocharger pressure ratio—the compressor operates with relatively low efficiency. This situation is further aggravated when the horsepower is held constant (line *D*). For this reason, with engines that must operate predominantly under conditions of variable rpm, it is desirable to design the turbocharger compressor so that it has the widest possible operating range, where good efficiency is achieved, even if some penalty in absolute peak efficiency must be paid. This can often be accomplished by the use of a vaneless diffuser, followed by a carefully designed scroll.

While in Fig. 4 the engine operating lines *A*, *B*, *C*, and *D* are shown superimposed on the *compressor* characteristic curves, it should be mentioned that the shapes and locations of these lines are not independent of the *turbine* efficiency and operating characteristics. Fortunately, however, it is much easier to design a turbine with an operating range sufficiently wide for turbocharger application than it is to design a wide-operating-range compressor. High turbine efficiency and a wide turbine operating range can most easily be achieved by using a centripetal-type turbine. The centripetal turbine has the additional advantage that it can readily be equipped with externally adjustable guide vanes, so that its flow characteristics can be changed during operation. By means of such guide vanes it is possible to radically change the shape of the engine operating line, as it appears superimposed on the compressor characteristic curves in Fig. 4. For example, the line of constant bmep and variable rpm can be so adjusted that the region of compressor instability is bypassed and the region of low compressor efficiency is avoided.

From the discussion of the matching of engine and turbocharger, and the engine operating lines (corresponding to different engine operating regimes), it will be clear that the matching of turbocharger to engine can be achieved in such a manner that the various torque characteristics illustrated in Fig. 5 are obtained. Line *A* of this figure shows the typical torque versus rpm curve of an atmospheric-aspirated engine. If this engine is turbocharged and the turbocharger so matched to it that optimum performance (that is, best fuel consumption and lowest exhaust temperature) is obtained at 100 per cent engine speed, the torque curve of the turbocharged engine will have the shape of curve *B*. If the best matching of turbocharger and engine is accomplished at a lower rpm, an increased low-speed torque (as shown by curve *C*) will be obtained. Finally, if the turbocharger is equipped with externally adjustable guide vanes, which permit suitably changing the turbine nozzle area with changing engine rpm, increased torque with decreasing engine rpm (such as shown by curve *D*) can be obtained. In the atmospheric-aspirated engine the maximum torque at low rpm is fixed by the smoke limit of operation. With externally adjustable guide vanes, the maximum torque at low rpm can also be fixed by the smoke limit, but it is more likely to be fixed by the point at which the compressor enters into its unstable range of operation. At any rate, it can be seen from Fig. 5 that, contrary to the belief often held by engine designers, turbocharging can be used to greatly increase low rpm engine torque.

As a matter of fact, turbocharging can bring about the ideal condition for which vehicle designers and users often ask; that is, constant brake horsepower output over a considerable range of engine speeds. This is made possible by the fact that the highly turbocharged engine operates with a very high combustion-air excess factor, so that at low rpm the combustion air excess factor can be reduced to increase the bmep sufficiently to hold the horsepower output constant, or very nearly constant. This is in sharp contrast to the atmospheric-aspirated engine, which usually operates with the minimum combustion-air excess factor at which it is possible to maintain smoke-free operation, thereby having no latitude for increasing the air-fuel ratio to make possible increased low rpm engine torque.

Another belief often held by engine designers not fully conversant with up-to-date turbocharger developments is that the ability of the turbocharged engine to accelerate is not as good as that of the atmospheric-aspirated engine. This is true only when the turbocharger is of the low-pressure type, has a poor combined efficiency, and/or has an unnecessarily high rotor wr^2. With high-pressure turbocharging and a high combined efficiency, the turbocharged engine normally operates with a much higher combustion air excess factor than the minimum corresponding to smoke-free operation. This permits the burning of large amounts of addi-

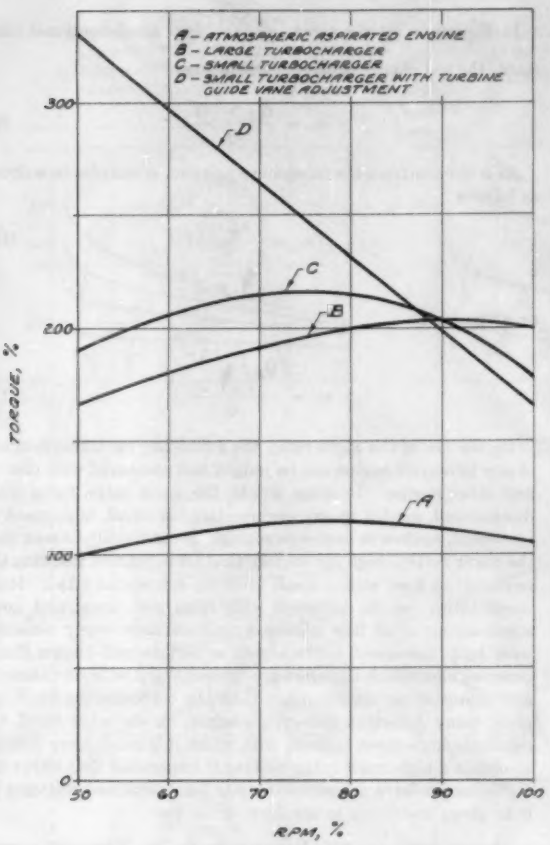


FIG. 5 ENGINE TORQUE CHARACTERISTICS

tional fuel during the acceleration period, resulting in an immediate response in the form of both engine and turbocharger acceleration. If the turbo wr^2 is low, this response usually is faster than can be obtained with the atmospheric-aspirated engine.

Constant-Pressure, Blowdown, and Pulse-Converter Systems. Since the exhaust connection between the engine and the turbocharger has an important bearing on the performance of the engine-turbocharger combination, a brief review of the several systems in use is indicated. Reference has already been made to the blowdown or pulse system, as distinguished from the constant-pressure system. In the latter, as the name implies, a virtually constant pressure is maintained in the exhaust manifold (which collects the exhaust gases from all cylinders in a single passage having a sufficiently large volume for plenum chamber effect, and conveys the exhaust gas flow to the single turbine inlet). In a constant-pressure system the energy in the exhaust pulsations does not become available to the turbine, which operates under conditions of steady flow and, preferably, with full admission. These operating conditions are ideal for the achievement of best possible turbine efficiency, which in many cases means that the losses incurred by sacrificing the pulse energy are counterbalanced to a large extent by the improved turbine performance. This is particularly true in the case of high-pressure turbocharging, because the higher the turbocharger pressure ratio the lower the pulse energy (in per cent of the total energy) that must be made available to the turbine.

In order to make available to the turbine some of the energy in the exhaust pulsations it is necessary to use a subdivided or branched exhaust manifold—in other words, to separately duct the exhaust gases from each cylinder, or from groups of cylinders (two, or not more than three in a group) having noninterfering exhaust phases. In the blowdown system, originated by the Swiss inventor Buchi, the separate ducts convey the exhaust gases into separate turbine nozzle box chambers, where the pressure is permitted to "blow down" during the end of the exhaust phase. In this system the turbine must operate with intermittent flow, only a limited segment of the stationary nozzles and the rotating blading being active at any one instant. This type of operation adversely affects the turbine efficiency, and for this reason the De Laval pulse converter system has been evolved, in which the outlets of the individual exhaust manifold branches are connected to the single inlet of the full-admission steady-flow turbine by means of an eductor device—the so-called "pulse converter"[2]—which is capable of converting the energy in the exhaust pulsations into a pressure rise. The result is that the turbine, while operating with steady flow, receives a preturbine pressure that is higher than the mean back pressure against which the engine must operate. It has been found by tests that the pulse converter system for recovery of pulsating exhaust energy is more efficient than the blowdown system for turbocharger pressure ratios higher than 1.8, whereas the blowdown system gives better performance for very low turbocharger pressure ratios of less than 1.5.

Even though in four-cycle engines under conditions of steady engine load and speed the gain to be derived by the utilization of the pulsating exhaust energy is small (or even negligible) if the turbocharger pressure ratio is high, utilization of this energy results in better acceleration of the turbocharger during load changes and in an improved performance of the engine-turbocharger combination at light engine loads.

TURBOCHARGING THE TWO-CYCLE ENGINE

The turbocharging of two-cycle engines represents the most recent and perhaps the most significant advance made in the development of the internal combustion engine since the advent of the high-pressure turbocharging of four-cycle engines. A number of test installations have shown that turbocharging can double, or very nearly double, the rating of the two-cycle engine, without any detrimental effect on its reliability and durability, because this large increase in output can be achieved without an appreciable increase in thermal loading. At the same time these test installations have shown that the fuel consumption of the two-cycle engine can be substantially improved by turbocharging—to the point where it is as good as that of the best four-cycle engines.

The turbocharging of two-cycle engines has been slow in arriving because it had to await the availability of turbochargers having higher combined efficiencies, and also because it presented a number of problems not involved in the turbocharging of four-cycle engines.

A two-cycle engine, in contrast to the four-cycle type, is not self-aspirating—it requires a separate source of air supply, capable of producing the necessary differential between intake and exhaust manifold pressures. A turbocharger by itself can only constitute this source of air supply if its combined efficiency is very high, because the energy available to the turbine from the exhaust of a two-cycle engine is small, and no assistance from the engine—such as that provided by the exhaust suction strokes of the four-cycle engine when the turbocharger efficiency is low—can be made available.

---

[2] Described in detail in SAE Paper No. 250, "New Developments in Turbocharging," by Rudolph Birmann; Jan., 1954.

Even with the highest possible combined efficiency, the turbocharger can only operate self-sustained and deliver, unassisted, the required air supply to the two-cycle engine (over the entire engine load range) if the air-handling characteristics of the engine are suitable. This means that the scavenging efficiency of the engine must be high, and the flow resistance must be low. In many cases, some means is also required to make a part of the energy in the exhaust pulsations available to the turbine.

A number of well-known European diesel engines in which the engine air is supplied exclusively by the turbocharger are performing satisfactorily in service. In all of these installations the prerequisites for successful operation are met; that is, the scavenging efficiency of the engine is high, the flow resistance low, the turbocharger combined efficiency is good, and provisions have been made for partial recovery of the energy in the exhaust pulsations.

In some of the typically American high-output two-cycle engines the flow resistance is high (chiefly because of the high engine rpm), and therefore the turbocharger (even if it has very excellent component efficiencies) cannot be self-sustaining, at least not over the entire engine load range. One solution for turbocharging these engines consists in retaining the crankshaft-driven blower and operating it in series or in parallel with the turbocharger.

Basic Considerations of Flow Through Two-Cycle Engine

Inasmuch as the flow resistance through the engine is one of the most important factors in the turbocharging of two-cycle engines, it is desirable—for analytical work—to have a nondimensional expression for measuring this flow resistance. From experience it was found that the most convenient measure is a nondimensional ratio which we shall call the *gas change process merit ratio*, derived as shown, using the following nomenclature:

Q_E = engine flow (referred to intake manifold conditions), cfm
Q = number of cylinders × rpm × (effective stroke × piston area + clearance volume), cfm
p = inlet manifold pressure, psia
Δp = pressure difference (inlet manifold pressure minus exhaust manifold pressure), psi
T_1 = inlet manifold temperature, deg R
λ_{sc} = scavenging air excess factor (ratio of Q_E to free air volume of fresh air in cylinders at beginning of compression stroke)
α = purity factor (ratio of fresh air in cylinders at beginning of compression stroke to Q)

From these definitions of λ_{sc} and α we see that

$$Q_E = \lambda_{sc}\alpha Q \quad [6]$$

From basic laws of flow

$$Q_E = K_E \sqrt{T_1} \sqrt{\frac{\Delta p}{p}} \quad [7]$$

where K_E is constant, regardless of turbocharger pressure level; however, only for a given engine. While dimensionally this constant is not an area, K_E is a measure of the flow resistance and can be considered the "equivalent orifice area" of a particular engine—in other words, the area of an orifice that presents the same resistance to the flow that is presented by the engine.

Combining Equations [6] and [7]

$$\lambda_{sc}\alpha = \frac{K_E \sqrt{T_1}}{Q} \sqrt{\frac{\Delta p}{p}} \quad [8]$$

In Equation [8] the term $\frac{K_E \sqrt{T_1}}{Q}$ is a nondimensional constant, the gas change process merit ratio m

$$m = \frac{K_E \sqrt{T_1}}{Q} \quad [9]$$

As is obvious from the foregoing equations, m can also be written as follows

$$m = \frac{\lambda_{sc}\alpha}{\sqrt{\frac{\Delta p}{p}}} \quad [10]$$

or, alternately

$$m = \frac{Q_E \Big/ \sqrt{\frac{\Delta p}{p}}}{Q} \quad [11]$$

By the use of the merit ratio, the suitability for turbocharging of any two-cycle engine can be judged and compared with that of any other engine. In other words, the merit ratio, being nondimensional, applies to any engine—large or small, high-speed or low-speed, uniflow or loop-scavenging. It can readily be seen that the merit ratio is high for engines that are capable of handling the required air flow with a small pressure differential (Δp). High merit ratios can be achieved with large port areas and good streamlining of all flow passages, and are most easily obtained with large low-speed engines such as certain well-known European engines, which therefore are turbocharged without the use of any source of air supply other than the turbocharger itself. A great many American two-cycle engines, on the other hand, are relatively high-speed engines, with which it is much more difficult to obtain a high merit ratio, making it imperative that either the turbocharger have an extraordinarily high combined efficiency or it be given assistance in one form or another.

Aerothermodynamic Relationships in Turbocharged Two-Cycle Engines

To illustrate the relationship between gas change process merit ratio, turbocharger pressure ratio, turbocharger combined efficiency, flow through the engine, various engine outputs corresponding to different exhaust temperatures, and various turbine nozzle guide vane settings, the curves in Fig. 6 are presented. These curves are the result of lengthy computations performed with the aid of the gas entropy chart, and are based on the following:

Ambient temperature, deg F 80
Intercooling, deg F ..... to 100
Turbine back pressure, psig 0
Turbine nozzle guide vane settings ..... *A*, *B*, and *C*
(where area B = 1.25 × area A, and area C = 1.50 × area A)

For the curves of Fig. 6(*a*) and (*b*) the turbine nozzle opening was fixed at guide vane setting *B*, and the turbocharger combined efficiency was assumed to be 60 per cent. Fig. 6(*a*) gives the specific air flow, in cfm per bhp, versus turbocharger pressure ratio, for four different merit ratios, and shows contour curves of bmep. The corresponding exhaust temperatures—for the same turbocharger pressure ratios and the same engine merit ratios—can be found in Fig. 6(*b*).

For Fig. 6(*c*) and (*d*) the merit ratio of the engine was fixed at 4.02, and for turbine nozzle guide vane setting *B* the specific air flow and exhaust temperatures are shown, as function of pressure ratio, for turbocharger combined efficiencies of 50, 60, and 70 per cent, with contour lines of engine bmep.

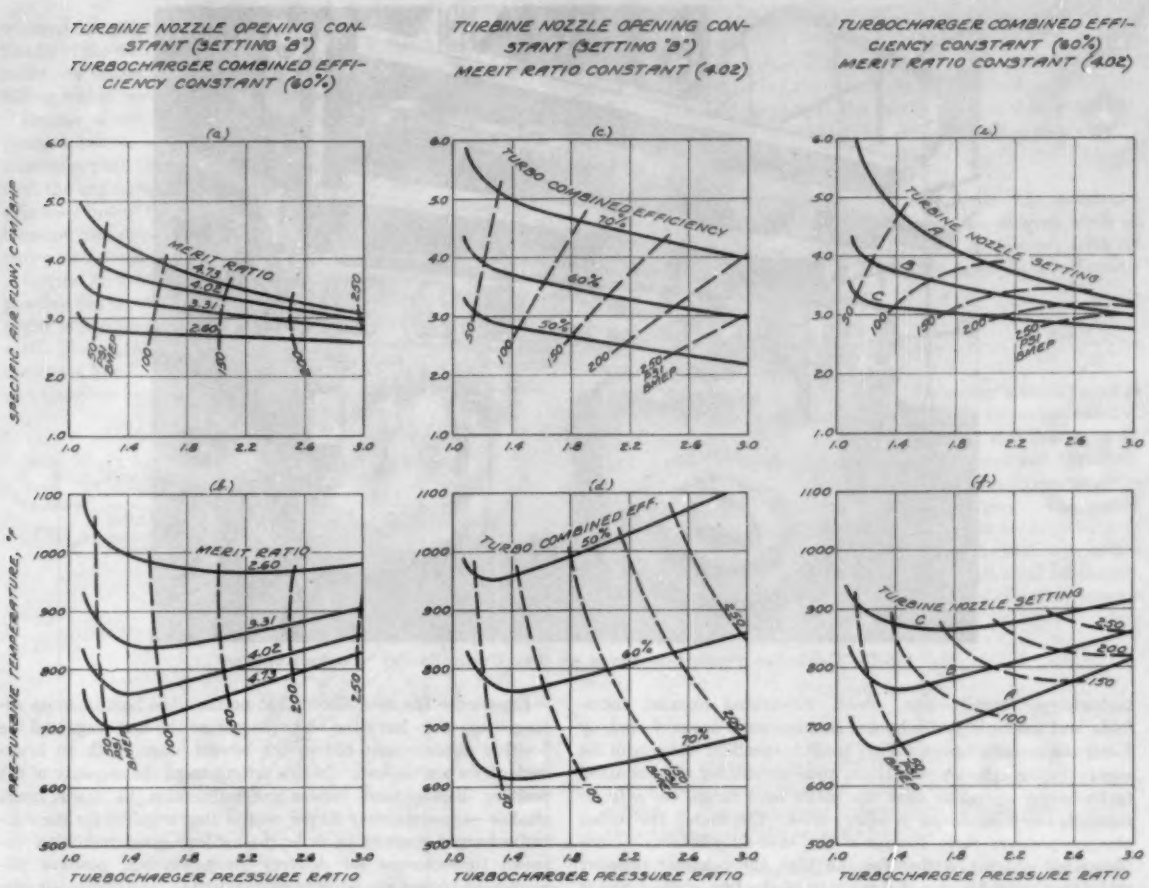

TURBINE NOZZLE OPENING CONSTANT (SETTING "B")
TURBOCHARGER COMBINED EFFICIENCY CONSTANT (60%)
TURBINE NOZZLE OPENING CONSTANT (SETTING "B")
MERIT RATIO CONSTANT (4.02)
TURBOCHARGER COMBINED EFFICIENCY CONSTANT (60%)
MERIT RATIO CONSTANT (4.02)
SPECIFIC AIR FLOW, CFM/HP
PRE-TURBINE TEMPERATURE, °F
TURBOCHARGER PRESSURE RATIO
MERIT RATIO
TURBO COMBINED EFFICIENCY
TURBINE NOZZLE SETTING

FIG. 6 AEROTHERMODYNAMIC RELATIONSHIPS IN TWO-CYCLE ENGINES

Finally, for a constant turbocharger combined efficiency of 60 per cent and a constant merit ratio of 4.02, Fig. 6(*e*) and (*f*) show the specific air flow and the preturbine temperatures, respectively, for guide vane openings *A*, *B*, and *C*, again with contour lines of engine bmep.

The curves of Fig. 6 answer a multitude of questions which continuously are raised in the turbocharging of two-cycle engines, and clearly bring out that engine gas change process merit ratio, engine bmep, engine temperature level, turbocharger combined efficiency, and turbocharger pressure level must have a very definite relationship to each other if self-sustained operation of the turbocharger is to be achieved. Since some of these relationships are not always clearly understood, and a discussion of all of the important facts that can be read out of Fig. 6 would be too lengthy, the following qualitative statements (of facts shown quantitatively by the curves) are offered.

It can be seen, for example, that in order to obtain high bmep with a two-cycle engine high turbocharger pressure ratios must be used. To obtain high bmep with a low temperature level—that is, low thermal loading—the engine must have a high merit ratio, which means a low flow resistance, and the turbocharger must have a high combined efficiency. The important effect of turbocharger combined efficiency on the exhaust temperature (preturbine temperature) is particularly apparent from Fig. 6(*d*). In this connection, it must be remembered that the exhaust temperature is an index of the thermal loading of the engine, and that high specific air flows, brought about by high turbocharger combined efficiency, lower the preturbine temperature and correspondingly reduce the thermal loading. By increasing the turbocharger pressure ratio (for any given bmep), the temperature level can be lowered and the thermal loading correspondingly reduced, as is best illustrated by Fig. 6(*f*). The curves in this figure also show the interesting effect of varying the turbine nozzle opening. The turbine nozzle guide vane opening can be changed for several purposes; for example, to increase the air flow through the engine and thereby lower the temperature level, or, at light engine load, to maintain the turbocharger rpm above the minimum self-sustaining speed.

In some engine-turbocharger combinations self-sustained operation of the turbocharger cannot be achieved—for example, if the merit ratio of the engine or the pressure level of the turbocharger is too low, or if the engine requires a high specific air flow for smoke-free operation, or the turbocharger combined efficiency is insufficient—in which cases aid, in one form or another, must be furnished to the turbocharger.

Utilization of Blowdown Energy. One of the simplest ways to provide the afore-mentioned aid is to make available to the turbocharger the energy in the exhaust pulsations. This can be done by the use of the blowdown system, involving subdivided or multibranched manifolds—preferably in combination with multiple

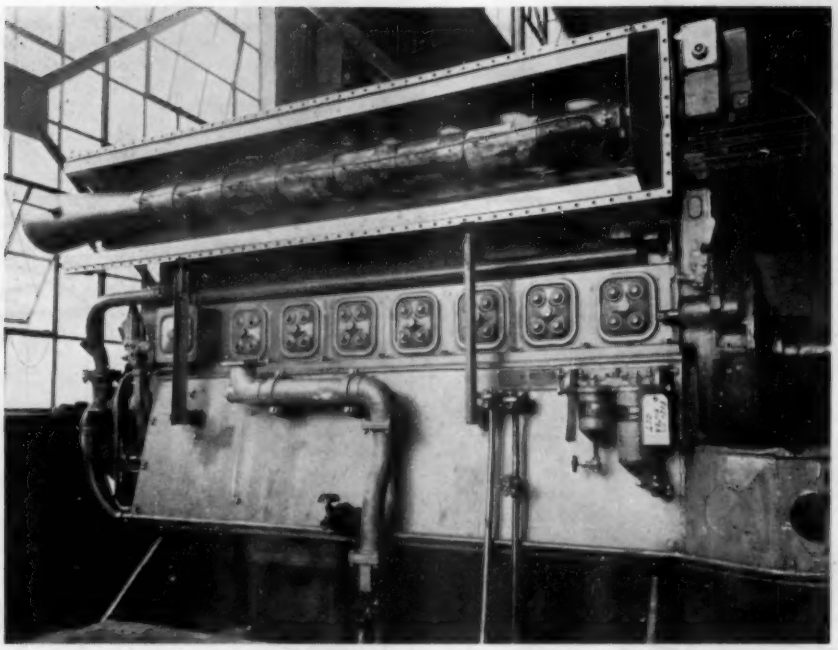


FIG. 7 DE LAVAL PULSE CONVERTER ON TWO-CYCLE ENGINE—COVER REMOVED

turbocharger installations. Such subdivided exhaust manifolds and multiple turbocharger arrangements are preferred by European manufacturers. They result in excellent recovery of the energy in the exhaust pulsations, thus permitting self-sustained turbocharger operation over the entire load range, *but only for relatively low turbocharger pressure ratios.* The higher the turbocharger pressure ratio the lower the over-all efficiency of the blowdown system, so that for the high turbocharger pressure ratios that must be used if the bmep of the two-cycle engine is to be increased to the point where this type is lighter and more compact than the highly turbocharged four-cycle engine, the blowdown system cannot be used. Therefore the author's company has evolved an alternate solution for the recovery of the energy in the exhaust pulsations, by adapting the pulse converter system to two-cycle engines, particularly those of the high-speed type. Such a pulse converter arrangement is illustrated in Fig. 7, which shows a test installation on a high-speed engine for which the pulse converter is particularly advantageous. The gain in over-all performance (in the case of high supercharger pressure ratio) through the use of the pulse converter system, as compared with the blowdown system, is brought about by the fact that the turbocharger operates with steady flow and with full 360-deg turbine admission. The action of the pulse converter permits the turbine to operate with a preturbine pressure substantially higher than the back pressure against which the engine must discharge. This difference in pressure represents the energy recovered from the exhaust pulsations.

Turbocharger and Positive-Displacement Blower in Series. Another form of aiding the turbocharger consists in retaining the crankshaft-driven blower of the two-cycle engine, and operating this blower in series or in parallel with the turbocharger.

Experience has shown that with the parallel arrangement the positive displacement blower can be very small, can have a very low power consumption, and need only be used during starting, idling, and very light engine load. At higher engine loads the blower can be disconnected.

Experience has also shown that perhaps the best in-series arrangement is to have the turbocharger as the first stage and the positive displacement blower the second stage, with an intercooler between the two. In this arrangement the capacity of the positive displacement blower generally must be made much smaller—approximately 60 per cent of that required for the nonturbocharged engine—in order to achieve good matching between turbocharger air delivery characteristics, positive displacement blower air delivery characteristics, and engine air consumption characteristics. With these components properly matched, the positive displacement blower absorbs engine shaft horsepower only in the very light engine load range, and becomes more and more "unloaded" with increasing engine output. Finally, as the engine load is increased, the crankshaft-driven blower acts as an expander and, as such, in addition to cooling the engine air supply, returns horsepower output to the engine crankshaft.

Operating Characteristics of Engine-Turbocharger Combination. The importance of high merit ratio and high turbocharger combined efficiency to the success of turbocharging the two-cycle engine have been pointed out. A third equally important requirement is good matching, which heretofore has been only briefly mentioned.

For the purpose of judging whether or not the engine air consumption characteristics and the turbocharger air delivery characteristics are properly matched, it is most convenient to plot the engine operating line, and superimpose it on the characteristic curves of the turbocharger compressor. This is illustrated in Fig. 8 which shows the operating lines of several turbocharger, intercooler, and engine combinations—one without the positive displacement blower and the others with it. The positions of all of these operating lines, relative to the characteristic curves of the turbocharger compressor, correspond to combinations that are well matched, at least for one operating point. It will be noted that for certain combinations of the various components this good matching is obtained over a wide range of engine operation,

whereas for other combinations portions of the engine operating line lie in the unstable range of the compressor and other portions in its low efficiency region, with only the intermediate portion falling within the good efficiency range.

Engine operating line *A* applies to the intercooled turbocharged engine without a positive displacement blower, and illustrates that the turbocharger and engine can be so matched that the engine operating line follows very closely the maximum efficiency line of the compressor. In other words, with this combination the compressor operates with high efficiency over the entire load range of the engine.

Engine operating line *B* represents turbocharger, intercooler, positive displacement blower, and engine, in series. If the intercooler is positioned between the engine and the positive displacement blower, and the turbocharger is the first stage, engine operating line *C* is obtained, which is steeper than either *A* or *B*, and therefore less desirable because at light engine load the turbo-

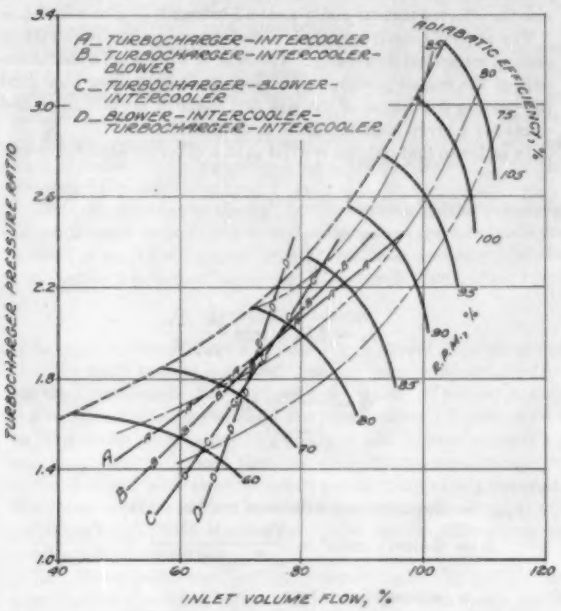


FIG. 8 TURBOCHARGER COMPRESSOR CHARACTERISTICS AND TWO-CYCLE ENGINE OPERATING LINES

charger compressor is forced to operate in the low-efficiency region. This situation is even more pronounced with the combination of positive displacement blower, intercooler, turbocharger, intercooler, and engine, the operating line of which is labeled *D* in Fig. 8.

SUMMARY

The diesel engine has been able to hold its own against the competition of other prime movers chiefly because it is capable of transforming the heat energy in the fuel into mechanical energy more efficiently than any of its competitors.

The future of the diesel engine looks bright, provided its thermal efficiency can be further improved, its output for a given investment in weight and manufacturing cost substantially increased without an impairment in its life and reliability, and its operating characteristics better adapted to various service requirements—all of which can be accomplished by high-pressure turbocharging, but will only be accomplished if the many aero-thermodynamic problems involved are clearly understood and adequately solved.

This paper was written with the idea of contributing to a better understanding of these problems, to offer practical methods for solving them, and to present the results of analytical investigations that have been carried out by the use of these methods.

ACKNOWLEDGMENT

The author wishes to gratefully acknowledge the assistance rendered by Mr. Anton A. Reinhart, and the diligent work of computation carried out by Mr. Stanley A. Lorenc, both of whom are associated with him in the De Laval Steam Turbine Company.

Discussion

P. D. HOBSON.[3] It is believed that the author's merit factor m should not be based on the total volume but on the swept volume. The clearance volume is frequently one of the first factors to be modified when a two-stroke engine is supercharged and, therefore, it should not appear in an expression intended to evaluate the initial suitability of an engine for supercharging. The use of effective stroke in the author's formula also calls for knowledge of a particular engine that might not readily be available to such a user as a turbocharger salesman looking for potential customers. Using the nominal swept volume and neglecting clearance volume the expression for m becomes

$$m = \frac{B \times 0.0283 \times T_1}{V \times \text{rpm} \times \sqrt{p\Delta p}}$$

where

B = blower output, cfm free air
V = engine displacement nominal (CSA × stroke) cu ft
0.0283 = ratio of normal temperature and pressure
Other terms as in the author's formula

The writer has calculated m for some 15 engines by the author's formula and from that just given and, though the absolute values of m so obtained are, of course, different, their relative values are little changed and can be used as a suitability guide in the intended manner.

This paper is a most valuable contribution to a rather neglected field.

R. H. MILLER.[4] The author is an exception among turbocharger specialists in that he has a broad knowledge and thorough understanding of the functions of the engine as well as the turbocharger.

Supercharging could very well be defined as the method of increasing the weight of air in the combustion chamber by multiple-stage compression, thereby producing a higher pressure and/or lower temperature than exists in the naturally aspirated engine.

Whether we deal with two or four-cycle engines, operating on oil or gas, we have only two basic systems by which turbocharging may be accomplished in a practical manner.

The oldest and most widely used is the Buchi system on which patents now have expired. It is characterized by a fixed cylinder compression ratio high enough for cold starting. The cylinder compression pressure, therefore, varies directly with the supercharging pressure.

[3] Research Associate, College of Engineering and Architecture Engineering Research, The Pennsylvania State University, University Park, Pa.

[4] Consulting Engineer, Nordberg Manufacturing Company, Milwaukee, Wis. Mem. ASME.

For example, a supercharging ratio of 1.6 results in a compression pressure of about 700 psi with a combustion pressure near 1000 psi. Supercharging the Buchi system to the highest pressure ratio shown in the paper, namely, 4 atm, results in a compression pressure of 1750 psi.

With intercooling to 100 F and a supercharging ratio of 1.6, a rating of 150 bmep may be carried with a normal thermal load such as we have with the normal Buchi system operating non-intercooled with 4 psi supercharging pressure at 120 bmep load.

Our second turbocharging system is new and radically different. It has been in commercial use only about four years but close to one-quarter million horsepower of the four-cycle type will be in operation at the end of this year. Its use is controlled by licensing under patents.

This system is characterized in diesel operation by a variable cylinder compression ratio which at starting and light load is the same as in the Buchi system but then decreases as the supercharging pressure increases to hold the compression pressure constant at the selected value.

For example, with a supercharging pressure of 3.1 atm at full load, the cylinder pressure ratio may be 15.5 giving a total first and second-stage pressure ratio of 48 with a resulting compression pressure of 700 psi. While this is equal to the compression pressure in the Buchi system with 1.6 atm supercharging pressure, the compression temperature is 250 deg F lower owing to the lower cylinder compression ratio and the weight of air in the combustion chamber is increased by about 20 per cent.

The combined effect of increasing the weight of the combustion air and decreasing the temperature at the point where heat is added by combustion raises the bmep to 200 psi when operating with the same thermal load and cylinder pressures as the Buchi supercharged engine at 1.6 pressure ratio and 150 bmep. That is, in this example, the variable compression-ratio cycle carries 33 per cent more load than can be carried with the Buchi system with the same cylinder pressures and the same thermal load.

Thus, in fact, the high ratios of which the author's turbochargers are capable require reduction of second-stage or cylinder compression ratios to hold cylinder pressures within practical limits.

In the variable compression-ratio system, the weight of air in the combustion chamber is increased by pressure as in the Buchi system but is further increased by lowering its temperature. It has been found that when supercharging, the mip is proportional to the absolute air-manifold pressure to the exponent 0.385 at equal combustion pressure and thermal load.

The intercooled rating curve in graph *E*, Fig. 1 of the paper, is in reasonable agreement with the result obtained with the writer's empirical equation which states, that mip at constant and equal thermal load is proportional to

$$\frac{P_2^{0.385} P_3^{0.275}}{T_2^{1.35}}$$

which, with constant compression temperature and P_3 rising with the supercharging pressure as in the case in graph *E*, makes mip proportional to supercharging pressure to the exponent 0.66. Using as a starting point the intersection of the intercooled curve at 1.5 ratio, the author shows 325 bmep at ratio 4, whereas a curve plotted from the same starting point with mip proportional to the 0.66 power of the pressure ratio on the abscissa and correcting from mip to bmep, the load at ratio 4 is found to be 285 bmep. The writer is of the opinion that the author's theoretically calculated rating curve is too high since the writer's empirical equation is derived from tests where the thermal loads were determined by measuring the heat flow to the water jackets.

The method is reported in a paper entitled "The Effect of Charge Temperature and Pressure on Rating of Diesel Engines and a Variable Compression Ratio Supercharging System for Two and Four-Cycle Internal Combustion Engines," presented at the Oil and Gas Power Division Conference in Dallas, Texas, 1951.

It would have enhanced the paper if the author had included in his two-cycle analyses the variable-compression-ratio system, which unlike the conventional systems he describes, can be operated with high supercharging pressures with correspondingly high mean effective pressures. Furthermore, owing to the 100 per cent longer scavenging period inherent in the variable compression-ratio system, the merit ratio is much greater than in the conventional system.

P. H. SCHWEITZER.[5] Much thought went into this paper and much useful information came out of it. The proper matching of engine and turbocharger is an important matter to which sufficient attention has not been given so far. Figs. 1, 2, and 6 of the paper should be very helpful in turbocharger design. The powerful effect of intercooling may surprise some engineers, and also the effect of turbine guide-vane adjustment, as shown in Fig. 5.

The author's merit ratio m is a useful concept but it should be defined somewhat differently. The m is, or should be, a characteristic of the engine porting or valving, an index telling how good the porting is because of the size and shape of ports, valves, and ducts for a given engine. Does it do that?

In order to find out, we rewrite m in a different algebraic form

$$m = \frac{\frac{Q_E}{Q}}{\sqrt{\frac{\Delta p}{p_1}}}$$

$$Q_E = L V_{disp} \text{ rpm } \frac{p_0}{p_1} \frac{T_1}{T_0}$$

$$Q = \frac{V_{disp}}{1 - \frac{1}{\epsilon}} \text{ rpm}$$

where

V_{disp} = displacement volume of engine, cu ft

L = delivery ratio[6] $= \dfrac{V_{del}(\text{cu ft NTP})}{V_{disp}(\text{cu ft})} = \dfrac{V_{del}(\text{lb})}{V_{disp}(\text{lb})}$

ϵ = compression ratio

p_1 = inlet manifold (air-box) pressure, psia

T_0 = 60 F = 520 K, P_0 = 14.7 psia

With these

$$m = \frac{L\left(1 - \frac{1}{\epsilon}\right)}{\sqrt{\frac{\Delta p}{p_1}}} \frac{14.7}{p_1} \frac{T_1}{520}$$

To see what is wrong with this, let us first visualize a positive-displacement blower, like a piston compressor with no leakage and no clearance volume. The compressor delivery is then equal to its displacement, say, 1.4 times engine displacement, and the cylinder porting is such that with 18.7 psia air-box pressure and 60 F air temperature the pressure drop across the cylinder will be 2 psi at 100 rpm (which is the rated engine rpm). Then

---

[5] Professor of Engineering Research, The Pennsylvania State University, University Park, Pa. Mem. ASME.

[6] "Scavenging of Two-Stroke-Cycle Diesel Engines," by P. H. Schweitzer, The Macmillan Company, New York, N. Y., 1949.

$$m = \frac{1.4}{\sqrt{\frac{2}{18.7}}}\left(1 - \frac{1}{\epsilon}\right)\frac{14.7}{18.7}\frac{T_1}{520}$$

$$m = 3.366\left(1 - \frac{1}{\epsilon}\right)\frac{T_1}{520}$$

If we do nothing but change the compression ratio of the engine from 14 to 13, using a thicker cylinder head gasket, the m goes down 0.6 per cent. With a blower of 80 per cent over-all efficiency the temperature rise in the blower is 21 deg F. With a less efficient (60 per cent) blower the temperature rise is 28 deg F and so the merit ratio goes up about 1 per cent. In other words with warmer intake air Q_E goes up but the lb/min flow (what the engine is interested in) goes down.

These changes are small; still why should the engine "merit ratio" go up when the blower efficiency goes down or when a head gasket is changed?

Another defect of the author's formula is that in a Kadenacy-type engine the merit ratio may become infinite or negative. In a blowerless Kadenacy engine,[7] the intake and exhaust pressures are equal, the pressure drop across the cylinder is zero, therefore m becomes infinite. In an engine with a small blower, with an exhaust turbine and with a pronounced Kadenacy effect fortified by an aspirator-type exhaust manifold, the measured exhaust back pressure can be higher than the air-box pressure. Then the merit ratio becomes negative.

Attention is called to the fact that the author's pulse converter is not different in principle from our aspirator-type manifolds described in an ASME paper[8] in 1952, and also previously at the International Internal Combustion Congress in Paris, May, 1951.

Author's Closure

In reply to Mr. Hobson's comments, it should be pointed out that the merit ratio m is derived from the basic equations governing the flow through a two-cycle type of engine. If engines having the same compression ratio and the same ratio of effective stroke to total stroke are compared by means of Mr. Hobson's simplified expression for m, values that are very nearly proportional to those obtained with the expression for the true m might result, but these values could not be used for a rigorous analysis of engines having different compression ratios and/or different ratios of effective to total stroke.

Mr. Miller points out that the bmep curve in Fig. 1(*E*) of the paper shows somewhat higher values than one that would result from his empirical formula, and he attributes this discrepancy to what he believes is a purely theoretical analysis which resulted in the curves of Fig. 1. It is more likely that this discrepancy is due to the fact that the curves presented in this paper are based on the conventional diesel cycle and do not take into account the effects of using the Miller cycle. With regard to Mr. Miller's belief that the curves have been theoretically calculated, it should be pointed out that while not each and every point on the curves has been confirmed by actual test results, the entire analysis resulting in these curves was based on heat-transfer coefficients and basic engine data that have been actually determined by tests, so the curves are truly representative of the performance of a typical non-Miller-cycle engine.

Professor Schweitzer suggests that the merit ratio should be an index telling how good is the porting, as well as the induction and exhaust systems, of a given two-cycle engine. Professor Schweitzer's criticism notwithstanding, the merit ratio is just exactly that. It should be and is affected by what Professor Schweitzer believes should not affect the true merit ratio. Specifically, the professor states that the clearance volume should not affect the merit ratio, and that this ratio should not be affected by the intake manifold temperature. Any true index, such as the merit ratio, must take exactly these factors into account and must refer to the total volume that must be scavenged, which includes the clearance volume. Similarly, the influence of the intake manifold temperature on the merit ratio is quite logical, as has been amply proved by exhaustive tests, conducted under the author's direction, for the Bureau of Ships, on a turbocharged General Motors 8-268A two-cycle engine.

With regard to Professor Schweitzer's discussion of the Kadenacy effect, he seems to completely overlook the fact that the engine is not concerned with the *total* exhaust pressure, but "feels," so to speak, only the *static* exhaust back pressure. If this fact is taken into account properly, it is obviously impossible to arrive at negative values of the author's merit ratio.

When Professor Schweitzer says that the De Laval pulse converter is not different in principle from his aspirator-type manifold (described in the references he has given, and disclosed in 1951 and 1952), he must mean that they both employ exhaust ducting, pipes, elbows, etc. Careful scrutiny of these references shows that the aspirator-type manifolds described and tested by Professor Schweitzer bear close resemblance to the manifold described in the author's patent No. 2,406,656, filed April 4, 1939, but do not contain the essential elements—primary nozzles, high-velocity momentum transfer passage, and diffuser. During the many years of experimental work that led to the design mentioned in the paper, it has been demonstrated over and over again that these elements are essential for good recovery of the energy in the exhaust pulsations, by transformation of this energy into a substantial pressure rise, which increases the Δp (difference between intake manifold pressure and exhaust back pressure across the engine) produced by the turbocharger.

[7] "The Kadenacy Effect," by P. H. Schweitzer, C. W. Van Overbeke and L. Manson, *Engineering*, vol. 160, September 28, 1945, pp. 241–244.

[8] "Research in Exhaust Manifolds," by P. H. Schweitzer, Trans. ASME, vol. 74, 1952, pp. 517–528.

Fuels for U. S. Navy Gas Turbines[1]

By H. F. KING[2] and H. V. NUTT[3]

The two principal fuels used by the Navy are Navy special boiler fuel oil and marine diesel fuel oil. These two fuels, one for high-pressure steam plants and one for high-duty diesel engines, supply the energy to propel practically all Navy ships and boats. At maximum efficiency, these plants can recover from 30 to 40 per cent of the energy input as shaft horsepower. Contrasted to this, a gas turbine can only recover up to about 25 per cent of the fuel energy. Therefore why would the Navy have an interest in a power plant which can recover such a small percentage of the Btu input? The answer is given in the paper.

Introduction

CONSIDER how Navy main-propulsion machinery is used. Based on a study of World War II operating data, full power was used only a fraction of 1 per cent of the time.[4] Further, more than 50 per cent of full power was used less than 2 per cent of the time. Thus the idea of using the gas turbine as a booster plant in conjunction with a base-load steam-power plant was born. The booster would operate only when the power required for the propulsion of the ship was in excess of that obtainable from the base-load plant, i.e., about 50 per cent of full power.

The Navy also is studying the space and weight requirements of the gas turbine. Presently, about 30 lb of machinery must be carried to generate 1 hp by high-pressure steam. On some of the newer diesel engines, the weight varies from 6 to 15 lb per hp. The weight of the booster-type gas turbine is less than 2 lb per hp. There is, therefore, a gain in weight advantage as high as 15 to 1 in the use of the gas turbine. As the booster turbine is expected to be operated for a small fraction of the total operating time, only a relatively small amount of fuel need be bunkered for its use. Thus the higher specific fuel consumption of the turbine plant has no appreciable effect in making the foregoing weight comparisons.

Navy Gas-Turbine Program

From the machinery-procurement and evaluation standpoint the Navy gas-turbine program is as follows:

Boeing 502 Gas Turbine. Over 20,000 operating hours have been accumulated on approximately 250 of these gas turbines installed in minesweepers for generator drive. Tests of a prototype minesweeping launch utilizing Boeing gas turbines for generator drive and propulsion are under way. Procurement of ten Boeing 502 gas turbines for installation in ten LCVP's has been initiated.

Two Boeing 502 gas turbines were evaluated as de-icers in the Arctic. By applying exhaust heat to metal equipment parts, the metal effectively dissipated heat and ice fell off equipment. This technique was a great deal more effective than trying to melt ice directly.

Solar T-45 Gas Turbine (45 hp). Several hundred of these gas turbines have been manufactured for a multitude of Defense Department applications. One turbine has been successfully subjected to a 1000-hr endurance test by the Corps of Engineers. This year, OPDEVFOR will evaluate two gas-turbine-driven fire pumps as de-icers in the Arctic. Exhaust gases will be used to melt ice.

Solar T-520 and T-522 Gas Turbines. The Solar T-400 turbine utilized for emergency 400-cycle generator drive is being replaced by an improved version (the Solar T-520). Four T-520 gas turbines also have been installed on minesweeping boats for generator drive.

The Solar T-522 gas turbine, a separate power-shaft version of the model T-520, has been developed and will be installed in a small high-speed boat for propulsion.

Metro-Vick G-2 Gas Turbine. It is considered desirable to test gas turbines under actual operating conditions in order to evaluate the problems associated with utilizing gas-turbine engines for propulsion as booster units. The developmental conversion of a PT boat at Philadelphia Naval Shipyard will incorporate prototypes of the first experimental gas turbine of this type to go to sea for the U. S. Navy. Delivery of turbines and conversion to gas-turbine-engine power is scheduled for 1955.

Rolls Royce RM-60 Gas Turbines. Main propulsion gas turbines, gear boxes, shafting, and propellers for experimental installation in the converted DE383 at Boston Naval Shipyard will be delivered in 1955. Conversion is scheduled for completion by the winter of 1955.

Modified Westinghouse J40. The first shop-test model of the marinized J 40 gas turbine has completed approximately 90 hr of test. Delivery of four completed engines to NBTL for test in a simulated shipboard installation of a Cosag plant was scheduled for April, 1955.

Packard Gas Turbine. A contract has been awarded to the Studebaker-Packard Corporation for the development of a turbine.

Project Wolverine. This development project provided for the design, construction, and testing of a semiclosed-cycle gas turbine for submarine propulsion. Testing as a surface-type power plant will be completed by approximately June, 1955.

Development of Liquid-Cooled Gas Turbine. Single-stage cold and hot-air tests have been run. Fabrication of a three-stage turbine is under way.

A more detailed description of many of these power plants has been published by Fowden, Peterson, and Sawyer in a paper entitled "The Gas Turbine as a Prime Mover on U. S. Naval Ships."[5]

Fuel Types

Gas-Turbine Fuel. Any plan of gas-turbine employment in the Navy would naturally include consideration of the type of fuel

---

[1] The opinions expressed and assertions made are those of the authors and are not to be construed as official or reflecting the views of the Navy Department or the Engineering Experiment Station.

[2] Head, Fuels and Lubricants Section, Bureau of Ships, Navy Department, Washington, D. C.

[3] Superintendent, Fuels and Lubricants Project, U. S. Naval Engineering Experiment Station, Annapolis, Md.

[4] "Prospects of Gas Turbines in Naval Applications," by R. T. Simpson and W. T. Sawyer, *Mechanical Engineering*, vol. 72, 1950, pp. 712–716.

Contributed by the Oil and Gas Power Division and presented at the Oil and Gas Power Conference, Washington, D. C., June 5–10, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, April 1, 1955. Paper No. 55—OGP-8.

[5] "The Gas Turbine as a Prime Mover on U. S. Naval Ships," by W. M. M. Fowden, Jr., R. R. Peterson, and J. W. Sawyer, ASME Paper No. 53—A-233.

to be burned. Thus, up to the present time, naval consideration of gas turbines has been on the basis of using either diesel fuel or Navy special fuel. Basically, all turbines under development are intended for the use of diesel fuel in the event no other fuel is suitable. Small high-speed gas turbines of the auxiliary type are applied in locations and operated in a manner which precludes the use of anything except a relatively volatile, nonviscous, easily burned fuel of the distillate type, similar to diesel fuel. This is also true of turbines for small craft propulsion. But in booster turbines used with steam-power plants, the possibility of using other than diesel fuel, i.e., residual fuels and lower grade distillates, is being explored extensively.

The U. S. Navy presently requires considerable quantities of diesel and Navy special fuels for shipboard applications. These fuels may be characterized as follows:

Diesel Fuel. Navy diesel fuel is essentially a straight-run fraction of petroleum. It may be characterized as belonging to the middle-distillate class of fuels. Typical values of properties are given in Table 1.

TABLE 1 PROPERTIES OF TYPICAL MARINE DIESEL FUEL

	Typical	Specification limit
Viscosity, SSU at 100 F	36	33–45
Cetane no	50	47 min
Sulphur content, % W	0.5	1.0 max
Composition	Straight run distillate	Straight run
Boiling range, deg F	400–650	725 max
Carbon residue, % W (on 10% bottoms)	0.1	0.2 max
Ash content, % W	0.0005	0.005 max

This type of fuel is used in all Naval diesel engines and in the small auxiliary gas turbines discussed previously.

Navy Special Fuel. Navy special fuel is a premium-grade residual product. It is made from the heavy residues (bunker oil) resulting from refining of crude petroleum which is cut back with a high percentage, up to 40 per cent, of lower viscosity distillate stocks in order to satisfy Navy requirements. This fuel is known world-wide. It is burned in all naval boilers and in a large number of naval auxiliary vessels used for military transportation. Although a residual-type fuel, the properties of this material are limited in order that ease of handling during naval operations is guaranteed. Typical properties of this fuel are given in Table 2.

TABLE 2 PROPERTIES OF TYPICAL NAVY SPECIAL FUEL

	Typical	Specification limit
Viscosity, SSU at 100 F	350	375 max[a]
Ash content, % W	0.05	1.0 max
Sulphur content, % W	2.5	None spec
Carbon residue, % W	10	15 max

[a] Estimated from actual requirement given at 122 F.

It will be noted that the carbon residue of Navy special is approximately 1000 times the carbon residue of Marine diesel fuel. The ash content is about 100 times that of diesel fuel.

Diesel Fuel in Auxiliary Turbines. To insure that no serious problems would be encountered in the application of diesel fuels to auxiliary-type gas turbines, an investigational program was initiated at the U. S. Naval Engineering Experiment Station, Annapolis, Md. The test setup is shown diagrammatically in Fig. 1. In this program, tests were made to determine whether a typical small, high-speed, open-cycle-type gas turbine would operate successfully with fuels covering the entire range of properties procurable under the Military Marine Diesel Fuel Specification MIL-F-16884 (SHIPS).

Installation for Testing Fuels. The gas turbine was installed in a test cell and coupled to a dynamometer. The high velocity of the exhaust gas was used in an eductor to ventilate the test cell. The warm test cell air was further heated as it passed over the outer periphery of the exhaust stack. A portion of this air was blended with outside air to regulate the temperature of the inlet air to the gas turbine. For convenient year-round operation, a controlled inlet temperature of 90 F was selected. The schematic diagram in Fig. 1 illustrates how the educted test cell air was blended with the outside air. This furnished a simple and economical method of maintaining a constant-temperature air supply to the gas turbine for the different fuel runs.

The fuels chosen ranged in volatility from those higher than obtainable in diesel fuel to a range lower than obtainable in normal diesel fuel. Their properties are given in Table 3.

The jet fuel was a JP-4 type procured from a supply at a Naval Air Station and represents service material. It was considerably more volatile throughout the boiling-point range than the other fuels tested.

The catalytic cracked distillate and marine diesel fuels were approximately equivalent in viscosity. The cracked fuel was appreciably aromatic in nature as indicated by its low smoke point, low cetane index, low aniline point, and low hydrogen-to-carbon ratio.

The heavy diesel was produced by blending a distillate from Middle East crude with a distillate lubricant stock to produce boiling points near the maximum limits allowed by the marine diesel specification. The heavy cracked distillate was a

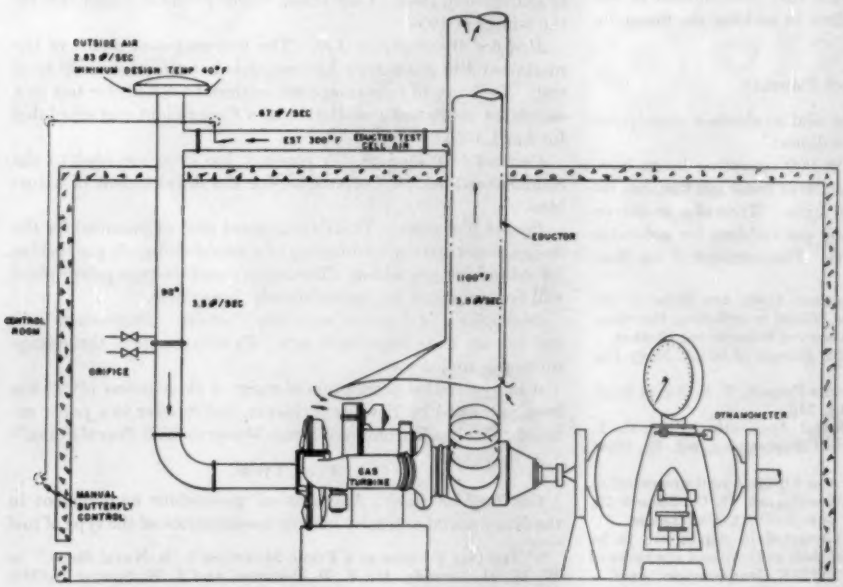


FIG. 1 GAS-TURBINE INSTALLATION FOR TESTING FUELS

TABLE 3 DISTILLATE-FUEL ANALYSIS RESULTS

	Jet-fuel JP-4	Diesel fuel[a]	Heavy diesel fuel[b]	Cracked distillate	Heavy cracked distillate
API degrees	48.4	36.8	35.8	26.2	18.2
Aniline point, deg F	125.4	157.5	165.7	102.8	95.3
Distillation, deg F					
Initial	154	369	400	343	284
10 per cent	240	449	473	485	457
50 per cent	327	522	549	539	580
90 per cent	464	597	680	596	723
End	528	646	712	625	
Viscosity SUS at 100 F	29.2[c]	36.0	39.4	36.7	44.6
Carbon residue, %, on 10% bottoms	0.05	0.08	0.03	0.19	12.0
Sulphur, per cent	0.08	0.10	0.80	0.42	1.81
Ash, per cent	None	None	None	None	None
H:C ratio	0.161	0.158	0.155	0.133	...
Naphtha insoluble, mg/100 ml[d]	...	127.4	87.1	104.1	10600
Smoke point, mm	22.1	20.8	19.8	8.0	6.6
Calculated cetane index	...	53	55	34	26
High heating value, Btu/lb	19904	19454	19441	18856	18249

[a] 0.1 per cent sulphur content.
[b] Mixture of 87 per cent Middle East diesel fuel (0.85 per cent S) with 13 per cent MS 2075 lube oil.
[c] Extrapolated.
[d] On residue from air-jet gum.

recycle stock normally used as a cutter stock for making Navy special fuel oil. It is a "wide-cut" material having front-end boiling points slightly lower than either of the diesel fuels listed in Table 3, but increasing rapidly so that about 75 per cent of the fuel boils at considerably higher temperatures than do the diesel fuels. It is also highly aromatic in nature.

It was found that combustion-chamber deposits were proportional to fuel boiling point and aromaticity. A jet fuel of the JP-4 type produced essentially clean combustors, Fig. 2. The moderately high boiling-point heavy-diesel fuel produced moderate amounts of soft, fluffy carbon in the combustors. A similar boiling-point fuel but more aromatic, produced by catalytic cracking produced even larger amounts of carbon which was hard and scaley.

The quantity of deposits produced by each fuel is shown in a relative manner in Table 4.

TABLE 4 RELATIVE DEPOSIT-FORMING TENDENCIES OF FUELS

	Per cent
Heavy cracked recycle stock[a]	100
Catalytic cracked distillate	45
Heavy marine diesel	12
Medium marine diesel	5
Jet fuel	2

[a] Liner buckled and turbine became inoperative after 5 hr of part-load operation.

Effect of Salt Water

During the course of the diesel-fuel evaluation, it became evident that dry diesel fuel would not always be available for use in auxiliary gas turbines. Therefore a program was initiated to determine the effect on turbine deposits and corrosion of salt water entrained in diesel fuel. A series of short tests using fuel containing up to 0.5 per cent of synthetic sea water demonstrated that sea-water contamination should be avoided. Early in this program, corrosion of ferrous parts in the fuel-handling system was noted, especially on springs and ball-check valves. The most serious problem encountered was that of freezing of the turbine rotor. Salts from the sea water combined with sulphur from the fuel to form high melting-point compounds which adhered to the turbine casing. During operation, the temperature in the combustion chamber was sufficiently high to maintain these depositions in a molten state. Thus no appreciable friction between stator and rotor was noted. After shutdown, however, these molten deposits solidified, bridging over the clearance between rotor and stator and freezing the rotor in place. This clearance space is 0.025 in.; thus only a small amount of deposit is required to form the bridge. The sticking was so severe that it was impossible to break the rotor loose by means of the starter fitted to the turbine. It was interesting that, in this series of runs, none exceeding 65 hr in length, noticeable corrosion of turbine-blade material occurred. The sodium content of the fuel due to the presence of the sea water was in the neighborhood of 65 ppm.

Distillate Fuel in Base-Load Turbines

The problems attendant to the burning of diesel fuel in large, base-load or booster-type turbines are less acute than in the

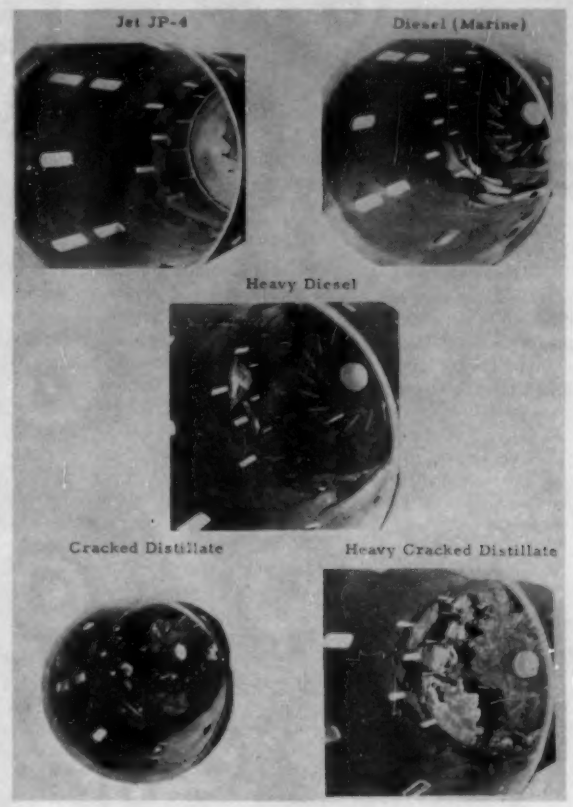


Fig. 2 View of Combustor Liner

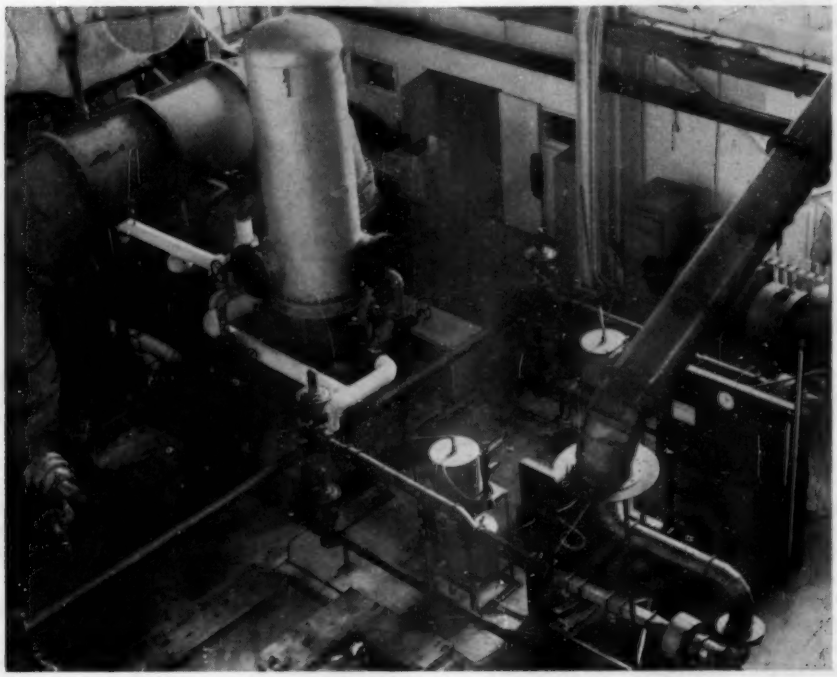


FIG. 3 INSTALLATION FOR TEST OF RESIDUAL-TYPE GAS-TURBINE FUELS

small auxiliary-type turbines. Considerable experience throughout the world using distillate-type fuels has been reported in the literature to prove this statement. Operation of a large base-load-type gas turbine at the U. S. Naval Engineering Experiment Station over a period of approximately 10 years, during which several thousands of hours of operation occurred, has also corroborated this view.

RESIDUAL FUELS FOR GAS TURBINES

The knowledge of the phenomena involved in burning residual fuels in gas turbines and the corrosion and deposits resulting therefrom appears to be well advanced. Reasonable success has been obtained by various petroleum companies and gas-turbine builders in applying residual fuel to gas turbines where the working gas temperatures do not exceed 1200 F. Various degrees of success have been achieved through several methods of attack, such as:

(*a*) Selection of the residual fuel to eliminate objectionable ash-forming constituents.

(*b*) The use of additives in the fuel to overcome deposits and corrosion otherwise present.

(*c*) The elimination of the undesirable metallic elements by water-washing or other fuel-treating methods.

(*d*) By controlled-combustion techniques wherein carbon particles of controlled size are formed which tend to prevent a build-up of corrosive deposits from the fuel ash on turbine parts.

Increasing the working gas temperature to values of 1400 and 1600 F tends to promote much greater difficulty in the successful use of residual fuels. It is contended by some that the only method of solving the corrosion and deposit problems at these higher temperatures is by use of fuel additives. Some members of this latter school appear to favor the use of additives containing magnesium. Other followers of this school have used silicon compounds advantageously in certain stationary installations. At one time oil-soluble calcium compunds were viewed with favor until additional experience indicated the superiority of other additives. Zinc and aluminum compounds also have shown some promise in combating serious corrosion of turbine blades.

Because of these apparent potential problems an exploratory program was instituted at the U. S. Naval Engineering Experiment Station to investigate the effects of burning Navy special fuel in turbines of the type normally classed as booster. Relatively high temperatures and high specific outputs were envisaged in such booster applications. Fig. 3 shows the laboratory test setup. The combustion chambers used are of a type designed for a commercial application. They are nominally 5 in. diam and burn 100 lb of fuel per hr. Air is charged to the combustors at a pressure of approximately 3 atm.

Metals of the type used in turbines presently under construction for the Navy were included in the investigation to determine the corrosive reaction of the fuel. These metal specimens exposed to the hot gases were actual turbine blades. They were placed in parallel downstream from the combustor. Through perfection of the geometry of the gas passages, a temperature variation of only about 50 deg F existed in the specimen compartment. In the investigations made, a Navy special fuel oil was used which was considered to be about midway between the best and poorest obtainable under the specification. The properties of this fuel are given in Table 5.

The same fuel was submitted to a number of laboratories for analysis. It will be noted from the range in the values obtained that techniques used in analyzing residual-type fuel apparently require considerable improvement in order to produce satisfactory reproducible results.

Figs. 4, 5, and 6 are photographs of metal specimens after 13, 28, and 98 hr exposure in the test rig at a nominal temperature of 1400 F. Deterioration of the specimen blades is clearly evident. Spectrographic analysis of the blade deposits revealed vanadium

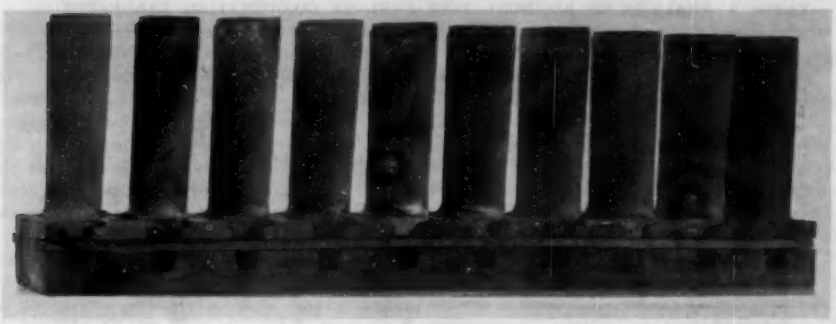


FIG. 4 GAS-TURBINE BLADES AFTER EXPOSURE OF 13 HR TO 1400 DEG F COMBUSTION GASES

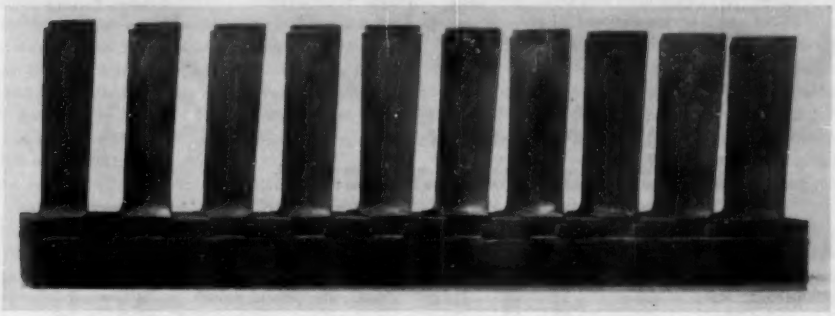


FIG. 5 GAS-TURBINE BLADES AFTER EXPOSURE OF 28 HR TO 1400 DEG F COMBUSTION GASES

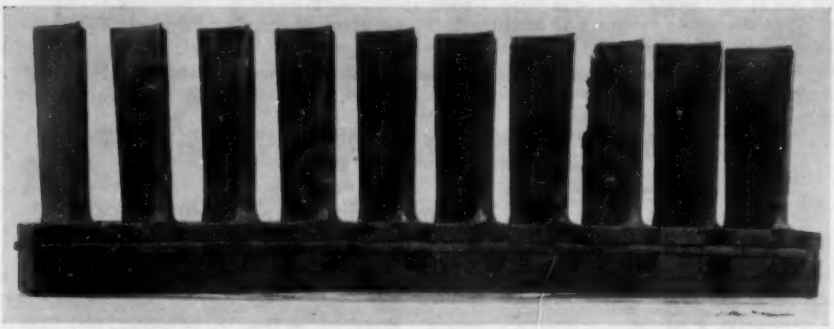


FIG. 6 GAS-TURBINE BLADES AFTER EXPOSURE OF 98 HR TO 1400 DEG F COMBUSTION GASES

to be the chief fuel-orginated metallic constituent and sodium second in order of quantity present. The metal weight loss of the specimens ranged from 0.6 to 16 per cent of the initial weight. Table 6 shows the composition of the metals tested and the weight losses after approximately 100 hr of operation at 1400 F. The two highest iron-content alloys suffered the greatest weight loss. The high nickel-content materials displayed the least weight loss.

Additional exploratory work to establish relative differences between marine diesel and Navy special fuels using AISI 310 (25-20) nickel-chrome steel specimens produced results shown in Table 7.

As a result of these exploratory tests, which tend to corroborate findings of others as published in the voluminous literature on the subject, there appeared to be no indication of a quick and clearly defined path leading to the goal of successful use of Navy special fuel oil in gas turbines. Because there is apparently no immediate answer to the problem, the Navy is seeking through research contracts to establish methods whereby successful gas-turbine operation on residual fuel can be accomplished. These projects will be described.

GULF OIL CORPORATION RESEARCH PROJECT

On July 1, 1953, the Gulf Oil Corporation was authorized to investigate the performance of residual-type fuels, burned at or near atmospheric pressures, for adaptability to gas-turbine use. The contractor is making up fuel blends varied so as to obtain extremes of fuel-ash constituents, sulphur content, and proportion of cracked and uncracked components in order to evaluate those fuel characteristics expected to be present in Navy special fuel oils. The contractor is exposing metal samples to ash deposition and corrosion and developing recommendations as to what can physically be done to Navy special fuel oil, with or without chemical additives, to make it a suitable fuel for gas-turbine application.

TABLE 5 ANALYSIS OF RESIDUAL TEST FUEL BY CO-OPERATING LABORATORIES

	No. of values	Test result		
		Low	High	Average
Gravity, °API	6	15.3	15.7	15.4[a]
Flash, P. M., deg F[b]	7	200	220	213
Fire, COC, deg F	5	285	290	289
Viscosity				
SU at 122 F	7	209	223	217
SF at 122 F	5	24.5	25.0	24.8
SU at 210 F	2	56.1	57.1	56.6
Pour point, deg F	7	Below −10[c]	0	−9
Water by dist., per cent	6	None	Trace	Trace
Sediment by extr., per cent	6	0.02	0.06	0.04
Water & sediment, per cent	7	0.05	0.10	0.08
Carbon residue, per cent	6	8.9	10.5	9.72
NBTL heater test	2	No. 1	No. 1	No. 1
Explosiveness	2	15	30	23
Ash, per cent	10	0.035	0.040	0.038
Water-soluble ash, per cent	1	..	..	0.01
Sulphur, per cent	3	1.62	1.67	1.65
Metals, ppm				
Vanadium (V)	10	109	145	126[d]
Sodium (Na)	8	8	29	15[e]
Nickel (Ni)	8	35	68	49
Iron (Fe)	4	17	38	26
Copper (Cu)	2	..	..	Present
Chromium (Cr)	1	..	..	Present
Silica and insoluble silicates	2	3	7	5
Aluminum (Al)	2	5	7	6
Calcium (Ca)	2	2	10	6
Potassium (K)	1	..	..	1
Zinc (Zn)	1	..	..	11
Magnesium (Mg)	1	..	..	Present
Lead (Pb)	1	..	..	Present

[a] One value of 16.4 ignored.
[b] Three COC results reported were 240, 250, 250 F.
[c] Averaged as −10 deg F.
[d] Includes a value of 115 calculated on assumption V_2O_5 was reported.
[e] Includes a value of 21 calculated on assumption Na_2O was reported.

TABLE 6 APPROXIMATE COMPOSITION AND WEIGHT[a] LOSS OF SPECIMENS

Metal specimen	Composition									Weight loss
	C	Cr	Ni	Co	Mo	W	Cb	Ti	Fe	
K-42-B	0.06	18	42	22				2.6	15	0.6
Nimonic	0.04	20	75					2.4	0.5	0.6
Inconel W	0.05	15	73					2.5	7.0	1.1
Inconel X	0.05	15	73				0.7		7.0	1.1
H.S. 23	0.40	24		63		6			2.0	1.7
H.S. 30	0.40	25	16	52	6				2.0	1.6
H.S. 31	0.50	25	10	55		7			2.0	1.8
Discaloy	0.03	13	26		3			2.3	55	15.7
N-155	0.35	20	20	20	3	2.5	1		34	2.9
AN-173	..	30	26	26		18			None	2.5

[a] Per cent of original weight after descaling at end of 98-hr run at 1400 F gas temperature. Origina weights ranged from 52 to 58 grams.

TABLE 7 EFFECT OF FUEL TYPE AND COMBUSTION GAS TEMPERATURE ON WEIGHT LOSS OF REFERENCE STEEL SPECIMENS AFTER 50-HR TEST RUNS

Mean gas temperature, deg F	Weight loss, mg per sq in. surface	
	Diesel fuel	Navy special
1450	..	320
1510	24	402
1555	33	535
1595	36	642

By subcontracting with the Westinghouse Electric Corporation, Gulf has obtained a static combustion burner which closely simulates conditions found in the Westinghouse J-40 gas turbine. Through technical co-operation with Westinghouse, this unit has been established at the Gulf Gas Turbine Laboratory in Philadelphia, Pa., and has been recently fitted for the initial tests.

Initial tests were conducted with diesel fuel to check out the performance of the combustion apparatus. Following this, two 100-hr tests were conducted using a Navy special fuel oil with an ash content of 0.04–0.06 per cent having a sodium and vanadium content of 10 ppm maximum. As was expected, no significant corrosion attack was experienced. The next fuel to be tested will contain 100–150 ppm sodium and 10 ppm maximum vanadium, and a third will follow having 10–20 ppm sodium and 100–150 ppm vanadium. With these fuels it is expected to obtain sufficient corrosion attack so studies can be started to counteract these phenomena with the use of chemical additives.

Ceramic Coatings

Studies of the resistance of ceramic coatings to oil-ash corrosion resulting from use of residual fuels have been, and are being, conducted at the U. S. Naval Boiler and Turbine Laboratory, Ohio State University, and the Stevens Institute of Technology. The investigation at the Naval Boiler and Turbine Laboratory was initiated because of corrosion attack of superheater support plates and seal plates in boilers aboard naval vessels. The corrosion apparently was occurring because of the slagging attack of the "fireside deposit" on the stainless-steel parts. The laboratory undertook a study to determine the resistance of existing ceramic, cermet, and metallic coatings to oil-ash corrosion. Initial tests consisted of immersing coated specimens in fireside deposit at elevated temperatures. In view of the fact that corrosion by the fireside deposit was very slow, efforts were made to develop an accelerated corrosion test. The accelerating action of carbon upon the slagging effect of vanadium and sodium attack, as indicated in a paper by B. O. Buckland,[6] indicated a method for obtaining an accelerated screening test. Use of carbon in a syn-

[6] "Residual Fuel Oil Ash Corrosion," by B. O. Buckland, C. M. Gardiner, and D. G. Sanders, ASME Paper No. 52—A-161.

thetic oil ash is considered logical because the presence of carbon particles as a result of incomplete fuel combustion during start-up, shutdown, or even constant operation of residual-fuel burners is very probable.

Refractory ceramic coatings, both government-developed and proprietary, which have been successfully employed for exhaust systems of aircraft engines operated on distillate fuels did not show much promise in resisting the slagging action of a synthetic oil ash consisting of 63 per cent sodium sulphate, 7 per cent vanadium pentoxide, and 30 per cent carbon when tested at 1700 F for 5 hr. Some measure of protection was assured in these immersion tests but the degree of protection was not considered sufficient to protect the basic metal during long periods of service. The behavior of these coatings indicated that the siliceous, or the glassy constituents, of these refractory coatings were too readily fluxed by the sodium and vanadium compounds. The results of these tests indicated that an oxide coating would be more desirable than the glassy ceramic coatings. Tests made on stainless-steel specimens coated with alumina afforded a marked degree of resistance to the slagging action of the synthetic mixture. However, the alumina was not sufficiently compatible with the thermal-expansion characteristics of the base and resulted in spalling of the coating upon cooling of the coated specimen.

The development at the Ohio State University of a flame-sprayable cermet coating, consisting of nickel and magnesium oxide for use in protection of aircraft parts against oxidation, offered a starting point for developing means of applying oxide-bearing coatings to metals. The nickel and magnesium-oxide coating did not successfully resist the corrosion or slagging attack of synthetic mixtures. Therefore the University is currently investigating other materials. Concurrently, the staff is attempting to determine the mechanism of corrosion of metals by the sodium and vanadium compounds. Three synthetic mixtures are being employed to determine the effects of high-sodium, high-vanadium fuels upon metals and refractory compounds. These mixtures consist of sodium, vanadium, and carbon as shown in Table 8.

TABLE 8 SYNTHETIC ASHES

Ash designation	Constituent		
	Na_2SO_4	V_2O_5	C
	Per cent by weight		
A	47.5	47.5	5
B	25	70	5
C	70	25	5

Promising materials virtually unattacked by synthetic mixtures at 1600 F are chromium metal, alumina, thoria, spinel, and two cermets—namely, alumina and chromium metal, and nickel and chromium borides. Attempts to coat metals with these materials are now under way.

Some successful experiments were performed at the Stevens Institute of Technology in conjunction with the Elracto Engineering Company, Hoboken, N. J., wherein powdered molybdenum specimens were successfully coated with cermet compositions. In view of these experiments, investigation involving the coating of powdered metals was sponsored by the Bureau of Ships. The Stevens Institute of Technology process consists of applying the coating as a "green" dry-pressed compact of powdered metal, sintering, and "hot coining" (forging in a die) immediately after sintering. It is believed that the coating applied by this method may be bonded to the basic metal mechanically as well as chemically. The objective of this investigation is to attain a multiple-coating system wherein the first coat is rich in metal and the final coat will be virtually pure oxide. It is believed that a multiple-coating system which will be compatible with the thermal-expansion characteristics of the basic metal can be developed. Specimens coated by Ohio State University and Stevens Institute will first be tested by immersing in synthetic ash at 1600 F for several hours. Promising coatings will then be subjected to testing in a microburner to determine the erosive-corrosive effects of the combustion products of residual fuel oils.

Oil-Ash Corrosion

Work on materials for gas turbines has been done with fuel-oil corrosion considerations in mind, but no blade alloy free of corrosion from the ashes of all the heavy fuels is envisaged. Alloy systems and certain metallic coatings have shown considerably improved resistance to these ashes, but they are not at present useful for turbine-blade applications.

The mechanism of oil-ash corrosion has been studied under Bureau of Ships sponsorship at M.I.T. The Bureau is now sponsoring a study at the University of Tennessee on the role of sodium in oil-ash corrosion.

Previous work has indicated that frequent cleaning of ash deposits is essential for freedom of corrosion from oil ashes.

Oil Ash Versus Oxidation Rate. The U. S. Naval Engineering Experiment Station has studied the oxygen absorption of deteriorated boiler tube-support plates. These supports were composed of 25–20 chromium-nickel steel and were coated with a tight adhering scale of which vanadium was the major metallic constituent and sodium one of the principal minor constituents. Presumably, the vanadium and sodium components were derived from the fuel burned under the boiler.

In this oxygen-consumption test, specimens of the supports, with the scale attached and also after descaling, were held at various temperatures in a closed oxygen system at atmospheric pressure for 3 hr. The gas volume consumed in the system was measured periodically. The results indicated that furnace-scale deposit adhering to the metal accelerated the consumption of oxygen in the system. Similar tests of the scale alone after its removal from the support plates did not produce oxygen-consumption rates of like magnitude. New 25-20 chromium-nickel steel after heating for 10 min in the powdered furnace scale at 1600 F also produced high oxygen-consumption rates. Temperature greatly accelerated the oxygen-consumption rate as shown in Table 9.

TABLE 9 OXYGEN CONSUMPTION VERSUS TEMPERATURE

Oxygen consumption, per cent	Temp, deg F
100	1600
30	1500
15	1400

These results indicate that scale containing vanadium and sodium may have an accelerating influence on the continued corrosion and deterioration of gas turbine-blade materials. It is planned to continue the oxygen-absorption study in that direction.

Cosag Fuel

It is highly desirable that gas turbines be developed without the very considerable disturbing influence of residual-type fuel. This could easily be accomplished by adopting a policy of developing gas turbines around the use of a distillate-type fuel similar to marine diesel fuel. However, there appears to be a very real shortage of such fuel in times of emergency. It appears that a large portion of military design is pointing toward the "middle of the barrel."

The Cosag power plants, described earlier, require either two fuels, one for steam generation and one for gas-turbine operation, or a single fuel meeting the requirements of both types of power plants. The fuel requirements of existing steam plants are well established and, in general, are fulfilled satisfactorily by Navy special fuel oil. New, higher output steam plants with higher heat-release rates and increased heat-exchange surface with at-

tendant closer tube spacing would be more effective if a fuel having a reduced amount of residual component was used. Gas turbines require fuels of low ash-depositing and corrosion tendencies, or essentially distillate-type fuels. The single fuel route to fulfillment of the requirements of both types of plants offers the following attractive features:

1 A fuel selected for its low deposit-forming and corrosion tendencies in gas-turbine application would be superior to Navy special as a boiler fuel with respect to both fireside deposits and corrosion in presently contemplated higher steam condition boilers and might be essential for boilers of any further increased steaming conditions.

2 The use of a single appropriate fuel would result in lower handling and maintenance costs and more effective operating practices.

The Bureau of Ships does not have a purchase specification for any fuel between diesel and Navy special. Therefore, as a basis for further discussion, Cosag fuel is defined as a fuel of low-ash and carbon-residue content and a reduced tendency from Navy special to form and retain salt-water emulsions. This fuel might be, for example, a heavy distillate comparable to commercial No. 4 fuel oil.

Aside from the decided benefits resulting from the use of a nonash fuel in the gas-turbine units of Cosag as compared to a residual-type fuel (Navy special) considerable advantage should also accrue in the burning of the nonash fuel in the steam power plant. Probable advantages would include:

(*a*) *Elimination of necessity for oil preheating which means:*

1 Absence of heater fouling.

2 Elimination of steam requirement for preheating.

3 Reduction in weight of power plant and complexity of piping.

(No down time for repair and cleaning of heaters.)

(*b*) *Greater rate of pumping and fuel transfer* at sea. The great majority of stocks furnished would be in the upper viscosity range of diesel fuel.

(*c*) *Prevention of boiler-tube slag* arising from fuel ash. This type is probably the most difficult to remove and more corrosive than that arising from sulphur and salt-water contamination.

(*d*) *Reduction of emulsifying problems* to a tolerable magnitude. Sea-water removal would then be accomplished with relative ease and tube deposits at high steaming rates would be minimized.

(*e*) *Use as emergency diesel fuel.* Viscosity requirements for a nonpreheating Cosag fuel together with the essentially nonash specification renders the fuel suitable on a strictly limited time basis as emergency diesel supply.

(*f*) *Reduced tendency to smoke* and form stack deposits and sparks. These characteristics are indicated by the relatively low carbon-residue property, about one-fifteenth of the maximum allowed in Navy special.

Application of Cosag Fuel. The immediate need for Cosag fuel is limited to experimental lots for evaluation in the Bureau of Ships' laboratories and for development of specification requirements. Near future needs would be for performance studies on the present Cosag test plant, on newly designed equipment, and possibly for full-scale fleet tests to compare Cosag with Navy special in existing equipment. Next in order would be a regular supply of Cosag for presently planned gas turbine installations and for future Cosag power plants.

Availability of Cosag Fuel. The development of an alternate fuel to residual fuel is not dictated entirely by machinery requirements. Availability of residual stocks is of real concern to those Navy planners who must think of a decade hence when planning and initiating today's research and development.

Residual fuel is the chief component of Navy special because it is cheap and available. It is usually blended with distillate fuel to obtain an acceptable viscosity. There is a continued trend toward decreased refinery yield of the low profit residual with attendant increase in lighter more valuable products such as gasoline and middle distillates. This trend is shown in Fig. 7. It is doubtless influenced considerably by the changing complexion of the petroleum-refining process in the United States. The thermal-cracking process is dwindling in relative importance. It is probable that some of the capacity now installed is in a state of disuse or partial use only. These trends result in a higher viscosity residue which requires ever-increasing proportions of distillate in the Navy special fuel blend. The ultimate situation will be one in which the small residual component downgrades (essentially contaminating) the higher-quality distillate component without realizing a substantially lower price or greater availability over the use of a wholly distillate fuel. This situation does not exist at present, of course, and distillate cannot be substituted for Navy special in an unlimited manner without severely handicapping other distillate applications.

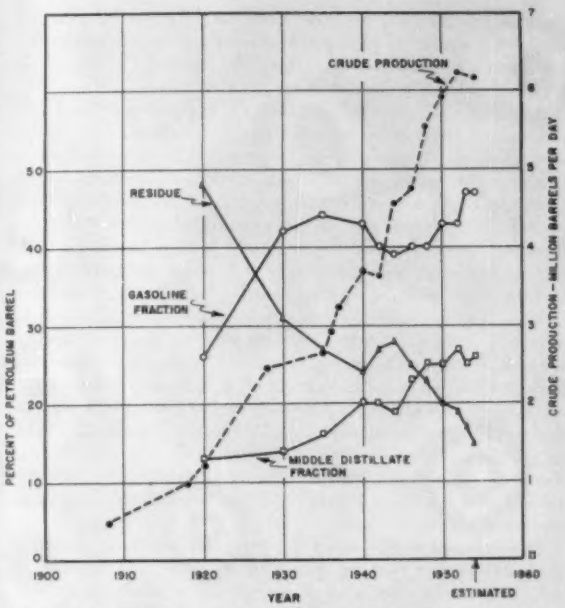


FIG. 7 CRUDE-PETROLEUM PRODUCTION AND PRINCIPAL CONSTITUENTS—U. S. ONLY

Recent military procurement of Navy special fuel included a five-million barrel contract quoted at $2.23 per 42-gal barrel based on bunker C costing $1.83 per barrel. Considering Cosag fuel to be equivalent in price to the cutter stock component of Navy special and further assuming bunker C would be cut back with 40 per cent of distillate to obtain suitable Navy special viscosities, then the price of Cosag fuel is estimated at $2.80 per barrel. Diesel fuel is currently being obtained at $3.34 per barrel in tanker shipments. Table 10 is a comparison of the procurement prices of distillate and residual products for the winter of 1954–1955.

TABLE 10 RELATIVE COSTS OF MARINE FUEL OIL

	Per gal	Per cent
Marine diesel	$0.080	100
Cosag fuel	0.067	84
Navy special	0.053	66
Bunker C	0.044	55

Based on the analysis in Table 10, the proposed Cosag fuel would

cost about one fourth more than Navy special, but only about four fifths of the cost of marine diesel.

New refining processes and new applications of refining processes, including "fluid coking," "contact coking," vacuum distillation, and deasphaltizing, are expected to promote the trend toward increased production of lighter products at the expense of residual-fuel yield. While the principal objective is usually the production of cracking-plant feed stocks for peacetime production of motor gasoline, these feed stocks could readily be diverted to Cosag fuel in the event of an emergency wherein rationing of civilian motor gasoline would be an essential conservation measure.

Specification for Cosag Fuel. A pilot specification has been developed for Cosag fuel as shown in Table 11.

TABLE 11 PILOT SPECIFICATION FOR COSAG FUEL

Gravity, °API	18–22	
Water and sediment, % by vol	0.1	max
Flash point, PMCC, deg F	150	min
Sulphur, per cent	2.0	max
Pour point, upper, deg F	20	max
Carbon residue, Ramsbottom, % by wt	1.0	max
Viscosity, SSU at 100 F	100	max
Ash, % wt	0.002	max[a]

[a] Up to 0.01 per cent acceptable if neither sodium nor vanadium exceeds 10 ppm.

No limitation is contemplated on the composition of the fuel. A blend of compatible distillates and residual types would be acceptable if conformance to the properties listed is obtained.

On the basis of combustion deposits and corrosion tests of experimental fuels, a specification will be prepared for use in procuring supplies for further laboratory studies, machinery development, and fleet trials.

Cosag Fuel-Development Program

Although a pilot specification, Table 11, has been suggested for Cosag fuel, several areas of doubt need to be resolved. To accomplish these ends, a number of fuels having properties in the neighborhood of those listed in the pilot specification are being procured and evaluated for performance characteristics. The properties of these test fuels are listed in Table 12.

The No. 1 fuel is a blend of No. 2 cracked distillate with Duo Sol extract from reduced crude. No. 3 fuel is a vacuum distillate from a U. S. continental crude source. No. 2 fuel is a blend of cracked gas oil and pressure distillate bottoms. No. 4 fuel is the same recycle material produced by catalytic cracking used in the distillate fuel tests described earlier. These four fuels indicate the many and varied methods which may be used to produce potential Cosag fuels. These fuels will be burned in a full-scale open-cycle auxiliary turbine of the type previously used for distillate-fuel studies. Ease of handling, sprayability, combustibility, corrosivity, and deposition characteristics will be measured.

Other combustion tests will be made using the full-scale, pressurized 5-in. static combustor test rigs shown in Fig. 3. In these tests a specimen compartment will be located in the way of the combustion gases. A group of metal specimens approximately $2 \times 4^1/_2 \times {}^1/_4$ in. will be exposed to the hot gases at temperatures ranging from 1200 up to 1700 F. The specimen materials have been chosen because of their potentialities as future Navy gas-turbine structural materials. The specimens and their compositions are listed in Table 13. AISI-310 alloy (25–20 chrome-nickel steel) has been selected as a standard for comparison purposes. Besides the fuels listed, their blends with residual fuels will be prepared and used to ascertain the effects of trace constituents, vanadium and sodium in particular, up to concentrations of 10 ppm.

The ever-present salt-water problem will also be given a thorough airing in the evaluation of Cosag fuels. Tight emulsions of residual fuel with water have defied standard water-separating techniques. It is hoped that Cosag fuel will allow development of satisfactory service techniques for water elimination.

Sulphur has been investigated many times in many places by many people. Limiting sulphur content to 1 per cent or less imposes severe availability handicaps. On the other hand, it would be difficult to obtain a Cosag type fuel having higher than 3 per cent sulphur. If availability considerations require the use of

TABLE 12 ANALYSIS OF COSAG TEST FUELS

	No. 1	No. 2	No. 3	No. 4
Gravity, °API	14.0	22.2	23.2	18.9
Flash, PMCC, deg F	210	290	340	
Pour point, deg F	−70	+45	+95	
Water and sediment, % vol	0.04	0.2	6.0	
Sulphur, % wt	1.93	0.46	0.26	1.81
Car. rc's., Ramsb., % wt	1.73	0.37	0.59	..
Cetane number	18.6	50.3	..	..
Cetane index	..	43.2	..	26
Ash, % wt	0.003	0.003	0.001	0.0–6
Viscosity, SU at 100 F	104	95	500	44.6
Aniline point, deg F	..	..	..	95.3
ASTM distillation:				
IBP deg F	396	586	512	284
10% rec. at deg F	467	642	659	457
30% rec. at deg F	489	657	711	..
50% rec. at deg F	536	672	..	580
70% rec. at deg F	650	694	..	..
80% rec. at deg F	665	712	..	..
90% rec. at deg F	..	732	..	723
End	..	..	..	766
Metals, approx ppm:				
Vanadium	1	1	1	
Sodium	3	2	1	
Calcium	3	1	1	
Magnesium	2	1	1	
Aluminum	2	3	1	
Silica	2	10	3	
Nickel	3	1	1	
Iron	5	3	1	
Zinc	..	1	.	

TABLE 13 COMPOSITION OF COSAG METAL TEST SPECIMENS

Alloy	C	Cr	Ni	Co	Mo	W	Cb	Ti	Fe	Al
Z	0.05	14	78						7	
Y	0.05	15	73				0.70	2.5	7	0.9
X	0.07	20	58	14	3.0			2.5	2	0.9
W	0.50	25	10	55		7			2	2
V	0.40	20	20	44	3.5	4	4.0		4	
T	0.35	20	20	20	3.0		1.00		32	
AISI-310	0.25	25	21						54	

fuels containing higher than 1 per cent sulphur, is there a benefit in restricting sulphur content at all? Apparently, gas density, gas temperature, and type and amount of sulphur compounds are important. These affect the dew point; i.e., the critical condition which affects the production of highly corrosive sulphuric acid. Are dew-point and acid concentration the only controlling factors and how are they related to fuel sulphur content? These are questions to be answered.

Residual Fuel in Base-Load Engines

Base-load gas turbines must have long life and high operating economy. Such gas turbines must operate at high inlet-gas temperatures to produce high specific output at high thermal efficiences. Great difficulties are encountered, however, due to the continuous exposure of metals to the high temperatures especially when burning cheap residual-type fuels. The high temperatures result in creep of metals and loss of strength; some residual fuels result in corrosion of blade materials and deposits.

Considerable benefit could be obtained by use of a compound, two-stage power plant wherein the temperature of the combustion gases is reduced in the primary stage to a value consistent with acceptable gas-turbine performance. The primary power plant would be specially designed so it could withstand high-temperature combustion and efficiently burn residual fuel. The combination free-piston gas producer-gas turbine engine[7] is suggested as a suitable compound power plant for development into a residual-fuel burner.

In this type of power plant, it is possible to obtain automatic adjustment of working fluid properties to obtain good part-load efficiencies and to make use of the ability to modify gas-flow rates to produce excellent torque characteristics. Also, some of the air compressed behind the power piston might be used to aid the atomization of the heavy fuel.

Some preliminary work has been initiated at the Engineering Experiment Station to determine how to use Navy special in a diesel engine. The results of this exploration should be of benefit in designing a composite, two-stage power plant for use with heavy fuels. Some of the potential problems which may require solution include:

(*a*) High cylinder pressures.
(*b*) Plugging of injectors.
(*c*) High cylinder and piston-ring wear.
(*d*) Piston-ring deposits (sticking).
(*e*) Valve-seat wear.
(*f*) Valve-stem deposits (sticking).
(*g*) Injector-valve sticking.
(*h*) Cold starting.
(*i*) Smoke.

Because the gas producer is still in its infancy, excellent opportunity is afforded for the exercise of ingenuity in designing the injection, combustion, and fuel-handling systems from the ground up so they will burn the cheaper grades of fuel most efficiently.

Acknowledgment

The authors are indebted to Mr. W. E. Robbins of the U. S. Naval Engineering Experiment Station for his valuable advice and help in the preparation of some of the tables and charts presented herein. Messrs. Frank Gessner, Henry Schab, and Jay Pasman also did yeoman service in obtaining the experimental data presented herein.

[7] "The Free-Piston Type of Gas-Turbine Power Plant and Applications," by Commander J. J. McMullen, USN, and R. P. Ramsey, Trans. ASME, vol. 76, 1954, pp. 15–29.

Discussion

A. M. Hudson.[8] The authors have considered their subject well, and have presented a good résumé of the present status of gas- turbine development and the fuels for these turbines. Additional emphasis might be placed on the question of how to compromise a fuel for use in both gas turbines and steam boilers.

Investigations at the Naval Boiler and Turbine Laboratory have been in progress for some time in an effort to determine the mechanism of oil-ash corrosion and slagging. Additives have been tried that range from metallic compounds to dolomite, including various organic compounds. Certain additives singly and in combination have been successful in reducing corrosion at various temperatures, but as yet there is no complete solution. It has developed that additives will modify corrosion tendencies, but not prevent them. An additive may change the composition of ash both physically and chemically but it will not gasify it or prevent its formation. A metallic coating on the boiler parts may serve to retard corrosion, but it will not prevent corrosion. And the higher the operating temperature the less the effect of these corrective measures becomes.

It appears then, that for boilers, there are problems that can be solved by using a fuel of better quality than the residual type. The authors show that Navy boiler fuel can be upgraded to be entirely distillate. Fuel quality should not be allowed to limit operating requirements; the petroleum industry can supply the needed fuels, but at an increased cost.

A question that must be answered is whether the increased cost of the higher-grade fuel can be absorbed by the savings in gas-turbine and boiler maintenance. An additional serious consideration is the time factor. Distillate fuels can be provided immediately to meet both gas-turbine and boiler requirements, but it may take some time to develop a satisfactory residual fuel for gas turbines.

D. M. Landis.[9] The authors have made a comprehensive study of the problem of utilization of various fuels in connection with gas-turbine operation by the U. S. Navy.

The writer cannot help but agree with the authors' statement that techniques in analyzing residual-type fuels definitely require considerable improvement in order to produce satisfactory reproducible results. Our experience has been that six different sets of figures will be obtained if one sample is divided into six parts and sent to six different laboratories for testing.

Also, the normal analysis of a residual fuel is far from complete. For determining the effects of centrifuging of residual fuels, BS&W content is of particular interest to us as is the per cent of water by distillation. Innumerable reports furnished to us by customers covering residual fuels, which were to be used as diesel-engine fuels, did not contain this information. Others did not list the per cent of ash by weight. In cases where residual fuels are to be utilized as either gas-turbine or diesel-engine fuels, it also would be advisable to have the ash analyzed for metallic components, such as vanadium, calcium, and sodium.

Our experience indicates that very little reliability can be placed on the normal fuel analysis that might be furnished to the user by an oil supplier. In most cases the analysis furnished was one made at the refinery and is not representative of the oil as received in the customer's plant. Normally uncleaned transport is used for residual fuel and considerable contamination is picked up by the oil in transit. Our present practice is to request a

[8] Lieutenant Commander, USN; Assistant to the Director for Petroleum, U. S. Naval Boiler and Turbine Laboratory, Naval Base, Philadelphia, Pa.

[9] Manager, Marine Division, The De Laval Separator Company, Poughkeepsie, N. Y. Mem. ASME.

sample of the oil as it is in the customer's tanks. This we analyze in our own laboratory before making any recommendation on purification equipment.

As reported in a recent paper,[10] there is a very successful gas-turbine installation in operation on residual fuels at the Central Vermont Public Service Corporation Plant at Rutland, Vt. This installation contains three 5000-kw General Electric gas turbine-generator sets which are operating at temperatures in excess of 1500 F on No. 6 residual fuel. The No. 6 fuels utilized at Rutland have had from 30–40 ppm of vanadium, 80–120 ppm sodium, and 150–250 ppm calcium. These fuels have been burned successfully in the gas turbines after a washing and centrifuging process.

When this installation was first started on No. 6 residual fuel, deposits formed very quickly on the first-stage nozzles and there was a noticeable loss of efficiency and power. In order to keep the power output up, it was necessary to remove this build-up of deposit by washing down to remove the water-soluble ash and by actually scraping to remove the majority of the ash, which was insoluble in water or steam. At that time the oil was being centrifuged merely for removal of suspended foreign material and evidently this was not the solution. By a large number of burner tests, General Electric determined that the main cause of both the deposition and corrosion was the presence of sodium and vanadium in the fuel. Chromium compounds were tried as additives for the sodium but increased slag formation resulted so this was stopped.

Magnesium sulphate was found to inhibit vanadium corrosion but ash deposits and corrosion due to sodium continued. Washing of the fuel was then tried for reduction of the sodium and calcium content and this worked out successfully. The oil was washed with water to transfer the water-soluble sodium and some of the calcium to the water and then the water solutions were removed from the oil by centrifuging.

With this system in operation, another problem soon developed, and this was the fact that some of the residual fuels had specific gravities just about equal to that of water. Normally in centrifuging, one liquid must be of higher specific gravity than the other or it is impossible to make a good separation. Centrifugal force will cause the heavier material to travel to the outside of the bowl where it can be drawn off from the lighter liquid, which moves toward the center of the bowl. Some of the No. 6 oils used at Rutland had specific gravities almost identical to that of water at the separating temperature, which made centrifugal separation impossible. To overcome this and to make separation possible, De Laval suggested that an inexpensive salt (calcium nitrate) be utilized as a wash solution and this would make the water phase heavier and simplify the separation of the oil and water. The calcium nitrate was later replaced with magnesium sulphate as the wash solution, as the calcium-nitrate solution gave too high a rate of deposition.

The washing system at Rutland normally reduces the sodium content of the oil down to less than 10 ppm and generally in the range of 4 to 7 ppm. In most cases, the calcium content is also reduced to the same range. Well over 3,000,000 gal of No. 6 fuel, which has been washed with either plain water or magnesium sulphate and then inhibited with an extra amount of magnesium-sulphate solution, has been burned in the Rutland gas turbines without any appreciable amount of deposition or signs of corrosion. The cost of washing this oil is less than $^1/_4$ cent per gal whereas the difference in cost between the No. 6 residual and No. 2 diesel fuel is approximately 4 cents per gal. Thus the operation is entirely economic.

A second washing installation for residual fuel for gas-turbine use is presently in operation at the Bangor Hydro Electric Company Plant at Bangor, Me. This system also is utilized in conjunction with a General Electric gas turbine.

Undoubtedly this same process of treating residual fuels could be utilized successfully with most all heavy fuels to reduce deposition and eliminate corrosion. It is also applicable to the treatment of fuels to be burned under high-pressure, high-temperature boilers where slagging is a major problem.

W. E. Young.[11] The authors are to be complimented on an excellent summary of the Navy's activities in the fields of gas-turbine and fuel testing. As a manufacturer of gas-turbine equipment, we are naturally vitally interested in such a program and in particular in the direction that Navy's thinking is taking as regards future fuel specifications.

Mention is made of a test in which diesel oil with sea-water contaminant was burned in a turbine with a resulting deposit which froze the rotor upon shutdown. It was implied that this deposit was sodium sulphate formed from a reaction of the chloride with the sulphur of the fuel. We would like to ask what the turbine inlet temperature was for this test. Pure sodium sulphate is molten at 1625 F which seems excessive for conventional turbines.

Synthetic ashes are very convenient for use in immersion tests but it is our opinion that the results obtained are frequently questionable. Ashes formed during normal combustion processes have been exposed to quite high temperatures. Decompositions may have occurred, but more important, reactions may have taken place resulting in compounds of different corrosivity than that of physical mixtures of the components. At times, we have noted a fusion of components taking place followed by reaction and eventual solidification even though the temperature of the mixture was gradually increasing. During the times when the mixture was liquid or reacting, measurable attack took place on immersed specimens. This condition might be rectified by preheating the synthetic ash in a controlled atmosphere until the desired reaction takes place followed by regrinding it to powder form.

The pilot specification for Cosag fuel should look very attractive to any gas-turbine builder. One thought occurs related to total ash and vanadium content. The Westinghouse residual-fuel-oil specification limits vanadium to 10 ppm and the total ash to 0.1 per cent which is not unreasonable. However, in this 0.1 per cent ash, there are almost certainly sufficient trace elements such as magnesium, silicon, aluminum, calcium, etc., to take care

[10] "Modified Residual Fuel for Gas Turbines," by B. O. Buckland and D. G. Sanders, Trans. ASME, vol. 77, 1955, pp. 1199–1209.

[11] Westinghouse Research Laboratories, E. Pittsburgh, Pa.

TABLE 14 ANALYSIS OF HEAVY CRACKED RECYCLE STOCK

Gravity, ° API	16.9
Sulphur, per cent	3.1
Aniline point, deg F	110
Flash, deg F (COC)	280
Viscosity, SSU @ 100 F	40
Pour, deg F	30
Heat content, Btu/gal	149000
Distillation:	
Initial boiling point, deg F	484
10 per cent	512
50 per cent	566
90 per cent	636
Final boiling point	670
Ash content, per cent	0.00035 (3.5 ppm)

Analysis of ash	Per cent on ash	Ppm on fuel
V as V_2O_5	0.13	0.005
Na as Na_2O	1.7	0.06
Ni as NiO	<0.1	<0.005
Fe as Fe_2O_3	9.5	0.33
Ca as CaO	0.91	0.03
Mg as MgO	0.88	0.03
Al as Al_2O_3	1.9	0.07
Pb as PbO (lead in shipping case)	29.0	1.0
Undetermined (SO_3, silica, etc.)	55.93	1.97
	100	3.50

of the existing vanadium. On the other hand, the Cosag fuel total ash is limited to 0.002 per cent or 20 ppm of which 10 ppm might be vanadium. We have found that to prevent corrosion an atomic ratio of magnesium to vanadium of at least three should be maintained and consequently some Cosag fuels might prove to be more corrosive than a straight residual oil. Granted, this is a rather remote possibility. The analyses given in Table 12 of the paper seem to be well on the safe side. Sodium, of course, further complicates the picture.

We have operated a full-scale combustor rig with no deposition or corrosion at temperatures of 1600 F using heavy cracked recycle stock (Table 14 of this discussion).

Based upon the foregoing test, heavy cracked recycle-stock fuels are satisfactory for gas turbines when atomized properly. The tests were conducted with no heating of the fuel oil, but with air atomization. When the combustion chamber is designed for the heavier distillate fuels, satisfactory combustion will be realized.

It would be interesting to know whether air atomization was used in the Navy tests of heavy recycle stock, and whether the combustion system was designed for a heavy fuel or light fuel.

Authors' Closure

The authors wish to express their gratitude and appreciation to Commander O. H. Landua for the wonderful way in which he briefed a rather lengthy paper and then so ably presented it. Our appreciation also goes to those who presented both written and oral discussions of the paper.

We are in complete agreement with Mr. Young's comments—even to our feeling that corrosion resulting from burning of Cosag fuel is "a rather remote possibility." In this respect, contamination of Cosag fuel with sea water appears to be the greatest cause for alarm. The recycle gas oil described by Mr. Young is similar to that used at the Engineering Experiment Station. Although containing negligible amounts of metallic compounds as supplied, the sodium was increased to serious proportions by adding sea water. A recent survey of boiler fuel stocks aboard naval vessels indicates many fuels are contaminated with 0.4 per cent sea water, or more. As indicated in the paper, the sodium content associated with this amount of sea water may cause serious and rapid corrosion and deposit build-up in the turbine. Consequently, Cosag fuel as finally developed should have characteristics which will allow it to be readily separated from water contamination with practical shipboard-type equipment.

The gas-turbine combustor used for burning the heavy cracked distillate at EES was designed for use with a light distillate such as Navy-type diesel fuel. The nozzle is a mechanical atomizing type, as distinct from the air atomizing type mentioned by Mr. Young. Consequently, the combustor is quite sensitive to fuel volatility, aromaticity, and carbon-residue content. A larger and longer combustor designed for heavy fuel and with air atomization burns this same fuel quite cleanly for extended periods of time with no combustion difficulties. We are, therefore, in accord with Mr. Young's observations that with proper combustor design burnability of Cosag fuel is no problem.

Although the deposit causing sticking of the turbine rotor in the salt-water contamination test was indicated to be sodium sulphate, this was merely the major compound present. About 30 per cent consisted of carbonaceous material and compounds of iron, calcium, magnesium, and silicon. Trace amounts of chromium, copper, cobalt, manganese, and molybdenum were also found. Melting-point tests of the deposit were made. Observations of the deposit as it was being heated are briefed as follows:

1300 F.....Particles (loose granules) of deposit displayed some tendency to stick together as in a sintering operation.
1400 F.....Original light gray color of particles grew much darker. Surface of particles became glazed in appearance and great tendency to stick together in continuous mass was displayed.
1500 F.....Molten mass with rough surface probably because complete melting had not occurred.
1600 F.....Molten mass with smooth surface.

The gas temperatures entering and leaving the rotor section of the turbine were about 1550 F and 1400 F, respectively.

It is quite probable that the deposit was a solid solution of several compounds with sodium sulphate predominating. If so, it is reasonable to expect one or more eutectic points to be displayed where melting-point temperatures lower than that of pure sodium sulphate would exist. The foregoing data support this theory and the "softening" or "sintering" temperatures indicated appear to be well within the maximum gas temperatures existing in the turbine.

Mr. Landis is certainly correct in stating that our paper is predicated on military necessity. Although we are not desirous of arguing the commercial applications of nonash (low sodium-vanadium) fuels, the backbone of the Cosag fuel specification and thus its outstanding characteristic, we do feel that such fuels have a real place in the future of the gas turbine in this country. The cost figures presented are based on actual procurement contracts and, of course as Mr. Landis assumed, represent large-quantity deliveries, i.e., tanker lots. Most of the Navy's fuel is obtained in this manner.

We are familiar with the operations at Rutland, Vt. Through the courtesy of the General Electric Company, Navy personnel have visited this installation several times. Although successful water washing is accomplished resulting in reduction of sodium content, such operations aboard a naval vessel appear to be impractical as sea water which contains large amounts of sodium would obviously be unsuitable as the wash water. It does appear, however, that the techniques described by Mr. Landis, which are evolving for use in stationary gas-turbine power plants, will be extremely helpful in allowing the gas turbine to compete economically with other sources of power.

Mr. Gessner has philosophized that wartime operation should require equipment developed to the point of burning satisfactorily the worst fuel obtainable. This is like hitching your wagon to a star. If one sets such a goal, recognized to be unobtainable, at least some steps will be made in the right direction so one is better off than if the attempt had never been made. Naturally, the best performance obtainable is desired, at any given time that is, from joint development of both fuel and equipment. This performance will always be superior to that obtainable if either the machinery or the fuel was developed separately without attention to the other.

Mr. Sullivan's staunch support of the thesis that availability of Cosag fuel will not be a stumbling block in its path is encouraging. Some remarks made to the authors by others are the source of the pessimism detected by Mr. Sullivan.

Mr. Buckland's point regarding the danger associated with salt water is well taken. We are definitely aware of it. Naval operations must always be aware of it. In fact, a large amount of our time is spent in voiding or inhibiting the detrimental effects of salt water. In this case it is hoped that the water-separating characteristics of Cosag fuel can be controlled by inclusion in the specification of a significant test method and suitable limiting values.

Effect of Size on the Inlet-System Dynamics in Four-Stroke, Single-Cylinder Engines

By D. H. TSAI,[1] WASHINGTON, D. C.

This paper describes a study of the problem of dynamic similarity in inlet systems of four-stroke, single-cylinder engines with long inlet pipes. A similarity rule formulated from a consideration of unsteady, isentropic, one-dimensional flow of a compressible fluid in the inlet pipe is checked by experiments.

Nomenclature

The following nomenclature is used in the paper:

A = dimensionless velocity of sound = c/c_0
A_p = piston area, sq in.
A_v = effective area of inlet valve, sq in.
A_2 = cross-section area of inlet pipe, sq in.
B = bore of cylinder, in.
C_p = specific heat at constant pressure, Btu/lb deg F
C_v = specific heat at constant volume, Btu/lb deg F
c = velocity of sound, ips or fps
D = diameter of inlet pipe, in.
e_v = volumetric efficiency defined by Equation [6]
e_{v0} = volumetric efficiency with zero inlet pipe
F = fuel-air ratio
F_s = stoichiometric fuel-air ratio
k = ratio of specific heats = C_p/C_v
L = combined length of inlet pipe and inlet port, in.
M_a = instantaneous mass rate of flow of air, lbm/sec
N = rpm
p = pressure, psia or in. Hga
R = gas constant, ft-lb/lb deg F
r = compression ratio
S = stroke of piston, in. or ft
s = piston speed = $2SN$, fpm or fps
T = temperature, deg F or deg F abs
t = time, sec
U = dimensionless velocity = u/c_0
u = velocity, ips
V_c = cylinder volume, cu in.
V_d = piston displacement volume, cu in.
X = dimensionless distance = x/L
x = distance from open end of inlet pipe, in.
Z = dimensionless time = c_0t/L
θ = crank angle, °ATC
ρ = density, lbm/in.[3]
Φ = velocity potential defined by Equation [1b]

Subscripts

0 = reference condition of atmosphere outside inlet pipe (in inlet tank)
1 = condition at open end of inlet pipe, immediately after entrance from atmosphere
2 = condition at cylinder end of inlet pipe (same as condition at inlet port)
avg = average value
c = condition in cylinder
e = exhaust condition

Other notations are defined as they appear in the text.

Introduction

The problem of dynamic similarity in the inlet systems of four-stroke, single-cylinder engines was brought out in a recent paper (1)[2] by Taylor, Livengood, and Tsai. In that paper a similarity rule applicable to engines of different sizes, and with different bore-to-stroke ratios, was formulated from theoretical considerations. The present paper proposes to examine this similarity rule in greater detail, and to evaluate the theory in the light of experimental data recently obtained at the Engines and Lubrication Laboratory at the National Bureau of Standards, Washington, D. C. The discussion will be limited to inlet systems with long inlet pipes. A comprehensive discussion of inlet systems without any inlet pipe may be found in reference (2).

Theoretical Considerations

General Equations for Inlet Process. A brief description of the general theoretical problem of inlet dynamics was given in Appendix A of reference (1). This section will review the important assumptions and equations pertinent to the problem of dynamic similarity. The details of the theoretical solution may be found in reference (3).

The problem of inlet dynamics may be formulated for the air[3] column in an inlet pipe of uniform diameter. The air may be assumed to be a perfect gas with $p = \rho RT$, and k = const. Further, the flow may be assumed to be isentropic and one-dimensional. Under these assumptions, the governing equation is a second-order partial differential equation of the hyperbolic type

$$(A^2 - U^2)\frac{\partial^2\Phi}{\partial X^2} - 2U\frac{\partial^2\Phi}{\partial X\,\partial Z} - \frac{\partial^2\Phi}{\partial Z^2} = 0 \quad \ldots\ldots [1]$$

where, as stated in the Nomenclature

$$A = \frac{c}{c_0},\ U = \frac{u}{c_0},\ X = \frac{x}{L},\ Z = \frac{c_0t}{L} \quad \ldots\ldots [1a]$$

Also

---

[1] Mechanical Engineer, Engines and Lubrication Section, Heat and Power Division, National Bureau of Standards, Washington, D. C.

Contributed by the Oil and Gas Power Division and presented at the Oil and Gas Power Conference, Washington, D. C., June 5–10, 1955, of The American Society of Mechanical Engineers.

Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, March 9, 1955. Paper No. 55—OGP-3.

[2] Numbers in parentheses refer to the Bibliography at the end of the paper.

[3] This is for engines using fuel injection into the cylinder. For engines using a premixed charge, the air will contain some vapor or gaseous fuel, but the equations derived here will not be changed.

$$U = \frac{\partial \Phi}{\partial X} \quad [1b]$$

The boundary conditions for this problem are as follows:

(*a*) *Open End of Inlet Pipe (Section 1).* If the flow is from the inlet pipe to the atmosphere, the pressure at section 1 may be assumed to be equal to the atmospheric pressure. This assumption gives the boundary condition

$$A_1 = 1 \quad [2]$$

If the flow is from the atmosphere to the inlet pipe, and if the entrance section is smooth and rounded, the flow process may be assumed to be quasi-steady and isentropic. Then the appropriate boundary condition may be written as

$$A_1^2 + \frac{k-1}{2} U_1^2 = 1 \quad [3]$$

(*b*) *Cylinder End of Inlet Pipe (Section 2).* If the inlet valve is closed, there can be no flow at section 2, and therefore

$$U_2 = 0 \quad [4]$$

If the inlet valve is open,[4] then air could flow either to or from the cylinder, depending on whether the pressure ratio $r_2 = p_e/p_2$ is smaller or larger than unity. For inflow to the cylinder, it may be assumed that the process across the inlet valve is quasi-steady, that from the inlet port to the inlet-valve section the flow is isentropic, and that there is no pressure recovery between the inlet-valve section and the cylinder. For outflow from the cylinder, similar assumptions may be employed. Under these assumptions, the correct boundary condition may be expressed as follows

$$d\left(\frac{p_e V_e}{p_0 V_d}\right) + (k-1)\left(\frac{p_e}{p_0}\right) d\left(\frac{V_e}{V_d}\right) = k\left(\frac{A_2 L}{A_p S}\right)\left(\frac{c_2}{c_0}\right)^{\frac{3k-1}{k-1}} (F \text{ or } G)\,\frac{d\theta}{6NL/c_0} \quad [5]$$

where

$$d\theta = \frac{6NL}{c_0}\,dZ \quad [5a]$$

and

$$F \text{ or } G = \left[1 + \frac{k-1}{2} M_2^2\right] M_2 \quad [5b]$$

but for F (inflow, $r_2 < 1$)

$$M_2 = \frac{u_2}{c_2} = \left[\frac{2}{k-1}\,\frac{1 - r_2^{\frac{k-1}{k}}}{\left(\frac{A_2}{A_v}\right)^2 \left(\frac{1}{r_2}\right)^{\frac{2}{k}} - 1}\right]^{\frac{1}{2}} \quad [5c]$$

and for G (outflow, $r_2 > 1$)

$$M_2 = \frac{u_2}{c_2} = -\left[\frac{2}{k-1}\,\frac{1 - \mathbf{R}}{\left(\frac{A_2}{A_v}\right)^2 \mathbf{R}^2 - 1}\right]^{\frac{1}{2}} \quad [5d]$$

with

$$\mathbf{R} = \frac{1}{2}\left(\frac{1}{r_2}\right)^{\frac{k-1}{k}} \times \left[1 + \left\{1 + 4\left(\frac{A_v}{A_2}\right)^2 r_2^{\frac{k-1}{k}} \left(r_2^{\frac{k-1}{k}} - 1\right)\right\}^{\frac{1}{2}}\right] \quad [5e]$$

It may be noted that when $r_2 = 1$, $F = G = 0$.

The situation involved here is shown in Fig. 1. Because of the periodic nature of the inlet process, it is known that the solution

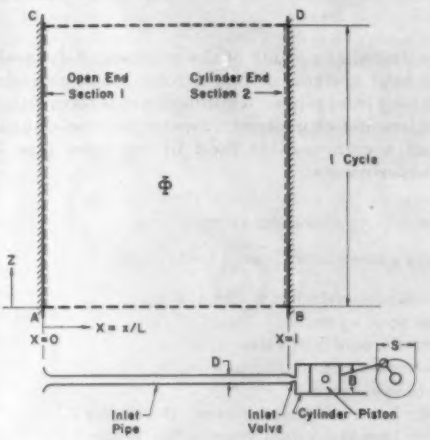


FIG. 1

must repeat itself after one complete engine cycle (two crankshaft revolutions). Therefore the problem is to find Φ, defined by Equation [1*b*], which will satisfy the boundary conditions of Equations [2] to [5*e*] along AC and BD, and in addition, the conditions along AB and CD must be identical in the interval of one complete engine cycle.

Volumetric Efficiency. The most important single quantity in this analysis is, of course, the volumetric efficiency of the engine, usually defined as

$$e_v = \frac{\text{Mass of air supplied to cylinder per cycle}}{\rho_0 V_d} \quad [6]$$

where ρ_0 is the reference inlet-air density, and V_d is the piston-displacement volume.

For computational purposes, it is also convenient to define an "instantaneous volumetric efficiency" as follows

$$\text{Instantaneous } e_v = \int_{\theta_{IO}}^{\theta} de_v = \frac{1}{\rho_0 V_d}\int_{\theta_{IO}}^{\theta} M_a\,dt \quad [7]$$

where M_a is the instantaneous mass rate of flow of air per unit time and θ_{IO} is the inlet-valve opening time (or crank angle). Integrated between θ_{IO} and θ_{IC}, the closing time (or crank angle) of the inlet valve, Equation [7] evidently reduces to Equation [6], and gives the volumetric efficiency of the cycle.

The term M_a in Equation [7] may be written as

$$M_a = \rho_v u_v A_v = \rho_2 u_2 A_2 \quad [8]$$

Here, by assumption, ρ_2 and u_2 are related to ρ_v and u_v through the area ratio A_v/A_2 and the isentropic relationship.

The relationship between Z, t, and θ is

$$dZ = \frac{c_0}{L}\,dt = \frac{c_0}{6NL}\,d\theta \quad [9]$$

[4] For convenience, the overlap of the inlet and exhaust valves will be assumed to be negligible. This should not affect the generality of the present discussion.

Also $$V_d = A_p S \quad [10]$$

Substituting these equations into Equation [7], one obtains, for the complete inlet process

$$e_v = \left(\frac{A_2 L}{A_p S}\right)\left(\frac{c_0}{6NL}\right)\int_{\theta_{IO}}^{\theta_{IC}} \frac{\rho_2}{\rho_0}\frac{u_2}{c_2}\,d\theta \quad [11]$$

Thus for a given A_2L/A_pS and $c_0/6NL$, the volumetric efficiency of the cycle depends on the density ratio ρ_2/ρ_0 and the Mach number u_2/c_0 at the inlet port.

Conditions for Dynamic Similarity. Since Equations [1] to [5e] involve only dimensionless parameters, it is clear that when these parameters are kept constant, the solution Φ to Equation [1] will be the same, regardless of the actual size of the engines. Under these conditions the flow pattern and the dynamic effect of the inlet system will be the same at the same engine crank angle. In particular, ρ_2/ρ_0 and u_2/c_0 at the inlet-port section will be the same during the inlet process, and therefore the volumetric efficiency, Equation [11], of these engines will be the same.

From Equations [5] and [5e], it is seen that the dimensionless parameters to be kept constant are the following:

Design Parameters

(*a*) A_2L/A_pS
(*b*) A_v/A_2 versus crank angle θ
(*c*) V_c/V_d versus crank angle θ

Operating Parameters

(*d*) Initial value of p_e/p_0
(*e*) NL/c_0

Engines having the foregoing design parameters in common may be called "similar engines." In these parameters, item (*c*) implies that the compression ratio and the ratio of crank radius to connecting-rod length must be the same. Item (*b*) implies that (i) the inlet-valve timing must be fixed, and (ii) the inlet cam profile and the area ratios between the inlet port, the inlet valve, and the combustion chamber in the cylinder must be similar. Consequently, A_v/A_p versus θ and A_2/A_p (or D/B), also must be kept constant. Combining the last condition of A_2/A_p = const with item (*a*), one observes that L/S must be kept constant separately. Therefore similar engines as defined here need not have the same bore-to-stroke ratio.

In the operating parameters, item (*d*) is largely controlled by the exhaust pressure p_e. Therefore p_e/p_0 must be kept constant. For similar engines, NL in item (*e*) is proportional to the piston speed $s = 2NS$. Therefore, at the same c_0 (or T_0), similar engines must operate at the same piston speed in order to have the same dynamic effect in the inlet system.

Departure From Dynamic Similarity. The foregoing conditions of dynamic similarity are based on theoretical considerations of thermodynamics and fluid mechanics. To the extent that the assumptions are valid, experimental confirmation of the theory reasonably may be expected. The assumptions, however, are not all valid, since the heat transferred to the fresh charge, and also to some extent, the friction between the fluid and the wall of the inlet pipe, are often not negligible. Moreover, in actual engines, the design parameters listed in the preceding section are seldom exactly similar. It is therefore of interest to examine briefly the effects of these factors on the similarity theory formulated in the foregoing.

(*a*) *Heat Transfer.* In references (1) and (3), it was found that assumption of adiabatic flow caused the computed volumetric efficiency to be consistently higher than the measured volumetric efficiency. However, since most of the heat transfer took place not in the inlet pipe, but at the inlet valve and in the cylinder, it was found that on a *relative* basis, using the volumetric efficiency for zero pipe, e_{v0}, as a reference, the computed and measured *relative* volumetric efficiencies, e_v/e_{v0}, were in satisfactory agreement (see Fig. 6 of reference 1). Also, in reference (2) it was shown that the effect of size on heat transfer in similar engines without inlet pipes was small. Therefore it would be reasonable to expect that in similar engines with inlet pipes, the *relative* volumetric efficiency would be the same under similar operating conditions.

(*b*) *Fluid Friction.* In reference (4), Jenny found that the frictional effect in a model of a pulsating engine-exhaust system (which was basically similar to the inlet system studied here) could be calculated with satisfactory accuracy by using an average value for the friction coefficient in the equation relating the friction forces at the pipe wall to the pressure drop per unit pipe length

$$\frac{dp}{dx} = \frac{\rho u^2}{2}\frac{4f}{D} \quad [12]$$

where ρ and u are, respectively, the local density and velocity of the fluid, D is the pipe diameter, and f, the friction coefficient, is defined in terms of the local shearing stress τ between the fluid and the wall of the pipe

$$f = \frac{\tau}{\rho u^2/2} \quad [13]$$

Jenny's results suggest that the effect of size on fluid friction in this type of problems could be estimated from steady-flow friction data. For a fluid in steady flow in a smooth pipe, f is a function of the Reynolds number (5)

$$\mathrm{Re} = \frac{\rho u D}{\mu} \quad [14]$$

where μ is the absolute viscosity of the fluid. In the ordinary range of engine operation, the average flow in the inlet system is usually turbulent (high Re), even in fairly small engines. Under these conditions f is usually small and does not vary too much with Re. Therefore the effect of friction in the inlet system ordinarily would be small, and would remain essentially unaffected by the size of the engine. On the other hand, in very small engines, the average flow may become laminar (low Re). In this case f is larger and increases more rapidly with decreasing Re. Thus the effect of friction probably would increase as the size of the engine is greatly reduced.

In the experimental data obtained here, there seemed to be some evidence of increased friction with reduced engine size (see discussion of Fig. 10). The results, however, were not conclusive and further study will be needed to verify the foregoing considerations.

(*c*) *Variation in A_v/A_p Versus θ.* The effective inlet-valve area A_v is usually small compared to the piston area A_p. Also, the kinetic energy of the fluid at the valve section A_v is usually dissipated in turbulence after the flow enters the cylinder, so that there is very little pressure recovery in the cylinder. This suggests that the effect of the size of A_p (or the shape of the combustion chamber) on A_v usually would be small. Thus it appears that the requirement of a fixed A_v/A_p versus θ (see discussion of item *b* in the preceding section) could be removed from the consideration of dynamic similarity.

With the restriction on A_v/A_p removed, A_2/A_p no longer has to remain fixed. Therefore L/S does not have to remain fixed independently, but rather A_2L/A_pS, or the ratio of the inlet-pipe volume to the piston displacement volume, must be kept constant. The conditions for dynamic similarity are then reduced simply to

the five conditions listed in the preceding section, with no auxiliary conditions attached.

It is interesting to consider here the special case of zero inlet pipe. The controlling parameters in this case are the same as those listed in the preceding section, except that parameters (a) and (e) must now be combined to eliminate the inlet-pipe length L, giving a new parameter $(A_2/A_p)(c_0/s)$ with A_2 now taken as the cross-section area of the inlet port. Combining this new parameter with parameter (b) of the preceding section, one obtains

$$(A_v/A_p)(c_0/s) \text{ versus } \theta$$

This is seen to be the reciprocal of the "gulp factor" defined by Equation [5] of reference (2)

$$\text{Gulp factor} = \frac{s}{c_0}\left(\frac{A_p}{A_v}\right)_{avg} \quad \ldots\ldots\ldots\ldots [15]$$

except that in Equation [15], A_v/A_p versus θ is replaced by $(A_v/A_p)_{avg}$[5] which is A_v/A_p averaged over the inlet-valve opening angle. Therefore, under the conditions of fixed parameters (b), (c), and (d) in the preceding section, engines without an inlet pipe must be operated at the same gulp factor in order to obtain the same dynamic effect and the same volumetric efficiency. This is of course the same conclusion reached in reference (2).

(d) *Variation in V_c/V_d Versus θ.* In a conventional engine, the effect of the ratio of crank radius to connecting-rod length on V_c/V_d versus θ is a higher-order effect. Also, in reference (1) it was found that a considerable variation in compression ratio (r = 6.0 to r = 10.3) had only a small effect on the dynamic characteristics of the inlet system. Thus it appears that the requirement of a fixed V_c/V_d versus θ, item (c) in the preceding section, also could be removed without serious effect on the similarity rule formulated here.

Experimental Setup and Procedure

Experimental Setup. The experimental data were obtained with a small Briggs and Stratton (B&S) engine modified for the purpose of the present investigation. To study the effect of size, these data were compared with the data of reference (1) obtained with a CFR engine. The physical characteristics of these engines are given in Table 1.

Fig. 2 gives one view of the general experimental setup. The inlet system consisted of a standard ASME orifice meter connected to two large surge tanks in series, and to the particular inlet pipe under test. In order to reduce flow loss, the transition section between the inlet tank and the inlet pipe, and that between the inlet pipe and the inlet port, were made smooth and well-rounded. The temperature and pressure in the inlet tank were respectively controlled by means of an electric heater and a throttle valve, both located between the two surge tanks.

Three different inlet cams were tested. These were designated as cams A, A′, and C. The timing of the valve events corresponded closely to those in the CFR engine, Table 1. The effective inlet-valve area A_v, under steady flow condition was obtained by the method described in reference (6). The results for both the B&S engine and the CFR engine are shown in Fig. 3. Fig. 4 shows that with cams A and C, A_v/A_p of the two engines were only approximately similar at various crank angles, although the $(A_v/A_p)_{avg}$ were nearly the same. With cam A', the lift was adjusted to make $(A_v/A_p)_{avg}$ one half of that of cam A.

Fuel from the city gas line was introduced to the inlet system at the throat section of a $\frac{1}{2}$-in. venturi located upstream of the inlet surge tank between the heater and the throttle valve. The amount introduced was controlled by a needle valve.

[5] This step is valid if the shape of the curve of A_v/A_p versus θ is not altered as A_v/A_p is changed.

TABLE 1 ENGINE AND INLET-SYSTEM SPECIFICATIONS OF B&S AND CFR (1) ENGINES

		B&S Engine	CFR Engine
Bore, in.		2-1/4	3-1/4
Stroke, in.		2-1/4	4-1/2
Compression Ratio		6.0	6.0
Ratio of Crank Radius to Connecting Rod Length		.233	.225
Method of Cooling		Head — Water-Cooled Cylinder — Air-Cooled	Water-Cooled
Type of Fuel		Natural Gas (Pre-mixed)	Gasoline (Cylinder Injection)
Inlet Port Length, in.		2-1/2	2-1/2
Inlet Port Diameter, in.		11/16	1-5/16
Valve Arrangement		L — Head	Overhead
Inlet Valve Diameter, in.		1.130 O.D. at Valve Head	1.35) O.D. at Valve Head
Inlet Valve Lift, in.		Cams A & C = .198 Cam A′ = .073	Cams A & C = .262
A_v/A_p vs. θ		See Fig. 4	See Fig. 4
Nominal Valve Events	E.O. °BBC	68	60
	E.C. °ATC	3	3
	I.O. °BTC	3	3
	I.C. °ABC	Cams A & A′ = 35 Cam C = 90	Cam A = 33 Cam C = 90
Inlet Pipe Length, in.	Cams A & C	Zero, L/S = 1.11 10.9, L/S = 5.97 16.7, L/S = 8.55 22.7, L/S = 11.2	Zero, L/S = .56 24.0, L/S = 5.97 36.0, L/S = 8.55 48.0, L/S = 11.2
	Cam A′	Zero, L/S = 1.11 24.0, L/S = 11.9 36.0, L/S = 17.1 48.0, L/S = 22.4	---
Inlet Pipe Diameter, in.		1.04, D/B = .46 .73, D/B = .33 .51, D/B = .23	1.50, D/B = .46 1.06, D/B = .33 .73, D/B = .23

The engine had a special water-cooled cylinder head for better control of the cylinder-head temperature. The cylinder block was cooled by a separate air blower. Lubrication of the engine was provided by an external pump which splashed a jet of oil on the moving parts. The oil temperature was kept constant by means of a heat exchanger.

The power output of the engine was absorbed by a small d-c electric dynamometer coupled directly to the engine.

Instrumentation. The orifice pressures and those of p_0 and p_e were measured with one-legged water or mercury manometers. The temperatures were measured with iron-constantan thermocouples connected to a Brown Electronik indicating potentiometer.

The air rate was measured with a standard 2-in. ASME orifice meter with flange taps across a 0.375-in-diam square-edged orifice (7). When the air rate was too low to be measured by this orifice meter, a special 1-in. orifice meter (not shown in Fig. 2) with an orifice diameter of 0.2004 in. was used. This meter was modeled after the standard ASME orifice meter, and was calibrated by means of a gasometer. The fuel rate was measured by another 1-in. orifice meter with an orifice diameter of 0.09995 in. It was calibrated with air, and converted for use with natural gas.

The large surge tanks were effective in eliminating pressure fluctuations and permitted accurate measurements of p_0, p_e, and the pressure drop in the orifice meters. The measured air rate and fuel rate were found to be reproducible to within 1 or 2 per cent.

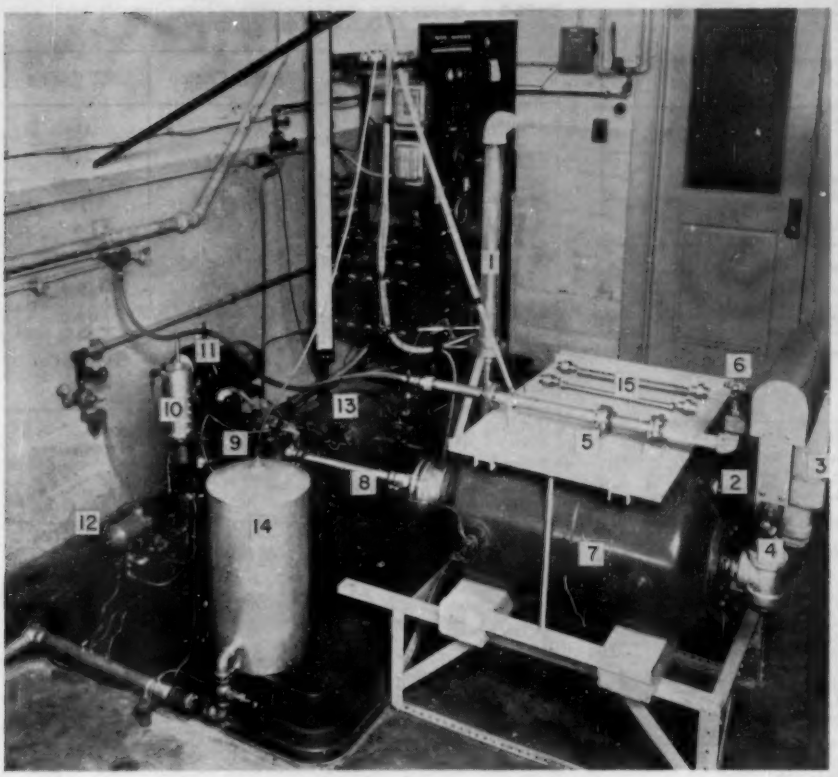


FIG. 2 EXPERIMENTAL SETUP

1 2-in. ASME orifice meter for air measurement
2 Surge tank
3 Electric heater
4 Throttle valve for controlling inlet pressure
5 1-in. orifice meter for fuel measurement
6 Needle valve for controlling fuel flow
7 Inlet tank
8 Inlet pipe
9 B&S engine
10 Cooling-water heat exchanger
11 Cooling air blower
12 Oil pump and oil heat exexchanger
13 D. C. electric dynamometer
14 Exhaust surge tank
15 Inlet pipe and adaptors

Engine speed was measured with an electric tachometer and a 60-cycle stroboscope. The brake torque was measured with an M.I.T. hydraulic scale (8).

Pressure at the cylinder end of the inlet pipe $^1/_2$ in. from the inlet port was measured with a strain-gage type of pressure pickup[6] for two series of tests. The pickup was calibrated under static condition. Because of the small output signal, the calibration probably involved some error. Nevertheless, the pressure records were found useful for qualitative comparison with the measured and computed inlet-port pressures in references (1) and (3), as in Fig. 5. In this figure the pressure waves at the inlet port are seen to agree well in amplitude and wave period under similar operating conditions.

Procedure. In Table 2 the operating conditions for the B&S engine are compared with those for the CFR engine (1). Here it is seen that the values of p_0 and p_e were kept the same as those for the CFR engine, as required by the conditions of similarity. The inlet velocity of sound, c_0, was also kept the same as in the CFR engine by adjusting the inlet temperature. (The properties of the fuel and the fuel-air mixture at $F = 0.067$ are summarized in Table 3.).

For cams A and C with the same $(A_v/A_p)_{ave}$ in both engines, the conditions for similarity further required that the two engines be operated at the same A_2/A_p (or D/B), L/S, and piston speed. Since the stroke of the B&S engine was equal to one half of that of the CFR engine, the B&S engine was run at twice the rpm of the CFR engine. The maximum speed of 4800 rpm, corresponding to 2400 rpm in the CFR engine, was limited by the dynamometer equipment.

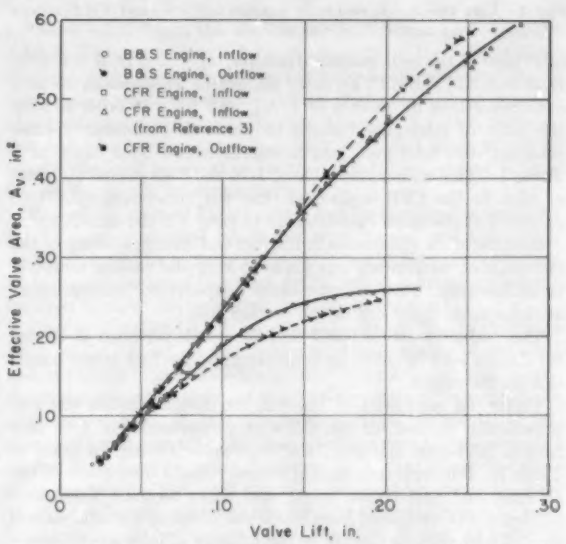


FIG. 3 EFFECTIVE AREA OF INLET VALVES OF B&S AND CFR ENGINES

For cam A′ with the $(A_v/A_p)_{ave}$ reduced to one half of its former value, the conditions for similarity required that A_2 also be reduced by one half. This was accomplished by changing to the

[6] Model P3B50A, Dynamic Instrument Co., Inc., Cambridge, Mass., 12-volt d-c excitation, without preamplification, connected to a DuMont Type 322A oscillograph.

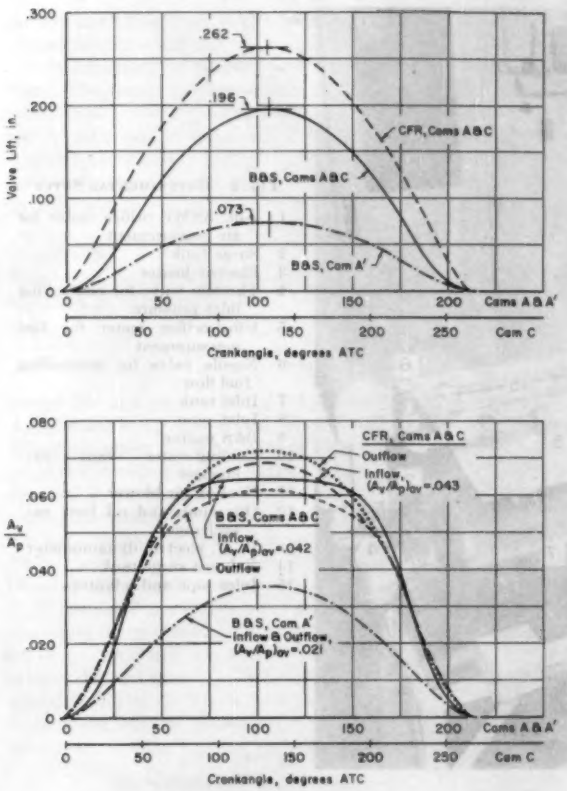


FIG. 4 LIFT AND A_v/A_p OF INLET VALVES OF B&S AND CFR ENGINES

TABLE 2 SUMMARY OF OPERATING CONDITIONS FOR B&S AND CFR (1) ENGINES

		B&S Engine	CFR Engine
Inlet Pressure, p_0, in.Hga		27.4	27.4
Exhaust Pressure, p_x, in.Hga		31.4	31.4
Inlet Velocity of Sound, c_0, fps		1160	1160
Inlet Temperature, T_0, deg F		100	100
Cylinder Head Water Temperature, deg F		210	180
Cylinder Block Temperature, deg F		155 (Fin)	180 (Jacket Water)
Oil Temperature, deg F		150	150
Fuel-Air Ratio		.067 (F_s=.060)	.080 (F_s=.067)
Spark Advance, deg BTC		30	Best Power
Engine Speed, rpm	Cams A & C	2000 to 4800 at 400 Increment	1000 to 3600 at 200 Increment
	Cam A'	1200 to 3600 at 200 Increment	---

TABLE 3 SUMMARY OF FUEL PROPERTIES

Composition, per cent by volume	
CH_4	88.0
C_2H_6	7.6
C_3H_8	2.4
C_4H_{10}	0.7
C_5H_{12}	0.3
N_2	1.0
Molecular weight	18.3
Stoichiometric fuel-air ratio, Fs	0.060
Fuel-air mixture with F = 0.067:	
Molecular weight	28.01
Gas constant, R, ft-lb/lb deg F	55.12
Cp, Btu/lb, deg F	0.269
Cv, Btu/lb, deg F	0.198
$k = Cp/Cv$	1.36

inlet pipe of the next smaller diameter, e.g., from $D/B = 0.46$ to $D/B = 0.33$ (Table 1), as these diameters were chosen to give area reductions by factors of 1, $1/2$, and $1/4$. In order to keep the ratio of inlet-pipe volume to piston-displacement volume constant, the inlet-pipe length was increased by a factor of 2 Table 1. This makes the value of L in the B&S engine the same as that in the CFR engine, so that the two engines must be operated at the same rpm in order to keep NL the same.

Because of the differences in the fuel and cooling systems of the two engines, no attempt was made to keep the various temperatures the same. The "cylinder-block temperature" was measured at the outside tip of one of the cooling fins.

The fuel-air ratio of 0.067 and the spark advance of 30-deg BTC were both found to be approximately the best power condition for the engine.

Under the conditions of Table 2, the dynamic effect was systematically studied for the different combinations of inlet-pipe length, inlet-pipe diameter, and inlet-valve timing, as listed in Table 1. For each run, careful measurements were made of the air rate, fuel rate, engine speed, and brake torque. Volumetric efficiency was computed from Equation [6] using the total mass of air and fuel and the density of the mixture at inlet conditions.

DISCUSSION OF RESULTS

Effect of Size With $(A_v/A_p)_{av}$ Kept Constant. The experimental results with cam A and cam C are summarized in Figs. 6 and 7, respectively, and compared with the corresponding results from the CFR engine in reference (1).

With cam A, Fig. 6 shows that the e_{v0} curve for the B&S engine was lower than the e_{v0} curve for the CFR engine, see also Fig. 8(a), but that the e_v/e_{v0} curves of the two engines were in close agreement for all the inlet pipes tested. This indicates that if the e_v curves were plotted for the various inlet pipes, the curves for the B&S engine also would be lower than those for the CFR engine. Fig. 8(b) shows two such curves for $L/S = 11.2$ and $D/B = 0.46$.

The third curve in Fig. 8(b) was obtained on the B&S engine with additional cooling of the cylinder block by means of a fine spray of water introduced to the cooling-air stream. Under these conditions, the e_v of the B&S engine was found to be in much better agreement with the e_v of the CFR engine. Thus, to a considerable extent, the discrepancy between the volumetric efficiencies of the two engines appears to be due to the effect of heat transfer being greater in the B&S engine than in the CFR engine. Other factors, such as the difference in the values of A_v/A_p versus θ, Fig. 4, in the thermodynamic properties of the fresh mixture, and in the method of introducing the fuel, etc., probably also contributed to the discrepancy between the two engines. In spite of all these differences, it is significant that the e_v/e_{v0} of the two engines agreed to within 2 or 3 per cent in most cases.

With cam C, Fig. 7 shows that both e_{v0} and e_v/e_{v0} continued to have the same trends of variation with s/c_0 in both engines, but the data differed considerably from those with cam A, Fig. 6, in many respects. Some of these are compared in detail in Fig. 8 for the case of zero pipe, and of $L/S = 11.2$, $D/B = 0.46$. For zero pipe, Fig. 8(c) (cam C) shows that the e_{v0} curve of the B&S engine was slightly lower than that of the CFR engine, but the difference was smaller than in Fig. 8(a) (cam A). For $L/S = 11.2$ and $D/B = 0.46$, Fig. 8(d) (cam C) shows that at low engine speeds the difference in the e_v curve was also smaller than in Fig. 8(b) (cam

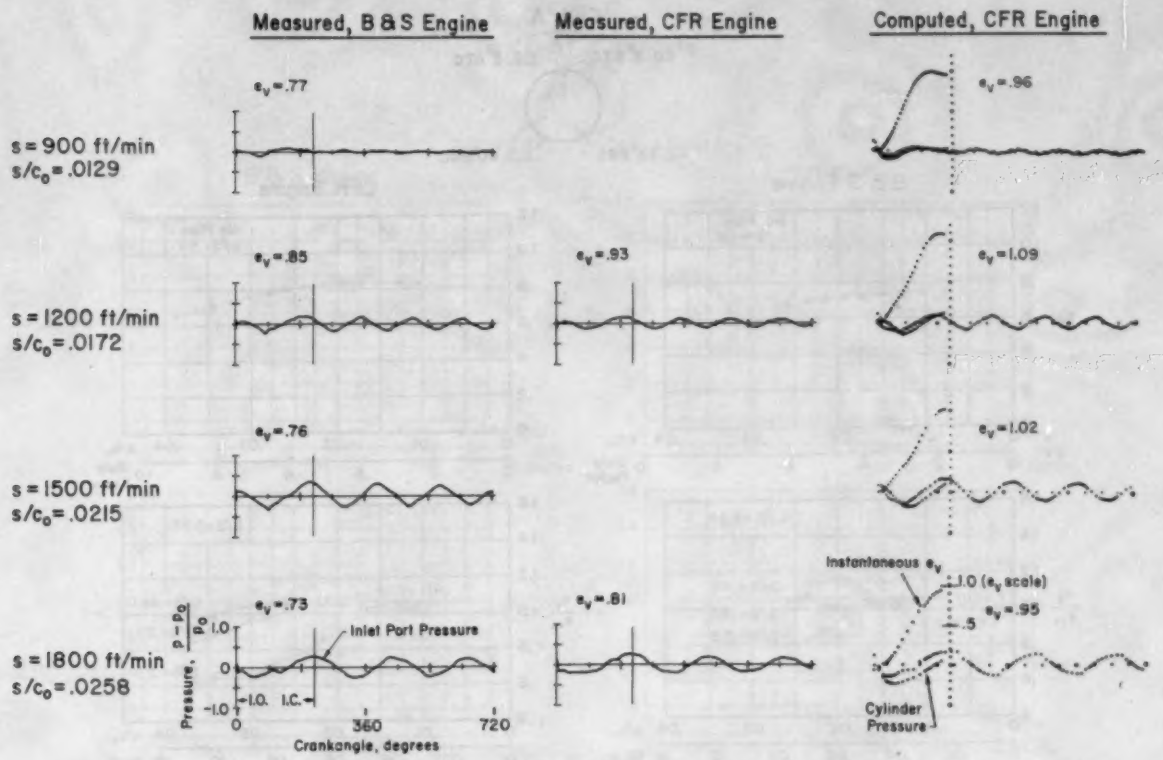


FIG. 5 MEASURED AND COMPUTED INLET-PORT PRESSURES OF B&S AND CFR (3) ENGINES WITH CAM A, $L/S = 11.2$, $D/B = 0.46$ (Other conditions as in Tables 1 and 2.)

A), but that at higher engine speeds, the difference was nearly of the same order of magnitude. Similar phenomena also could be observed in the other cases.

The variations between the two cams appeared to be caused by the dissimilarity in the flow characteristics of the inlet valves of the two engines, Fig. 4. In both engines, the values of A_v/A_p for inflow and outflow were nearly the same at low valve lifts, but at higher lifts, A_v/A_p for inflow was higher than that for outflow in the B&S engine, and the reverse was true in the CFR engine.

The probable effect of this difference may be seen from Fig. 9 taken from reference (3). This figure shows a series of eight computed solutions for a CFR engine under the conditions of Tables 1 and 2. These solutions cover the extreme range of inlet-pipe length and engine speed investigated here. The curves of interest are the instantaneous e_v curves which show the quantity of air in the cylinder during the inlet process. A rising curve indicates that air is flowing into the cylinder, and a falling curve indicates that air is flowing out from the cylinder. The end point of the curve, i.e., the intersection of the curve with the vertical axis located at the closing angle of the inlet valve, gives the volumetric efficiency of the cycle.

From this figure it is seen that very little outflow occurred in all the four cases with cam A. With cam C, the outflow was also quite small for $L/S = 11.2$ and $D/B = 0.46$ at 1800 fpm piston speed; but in the other three cases, the outflow was much more pronounced. Moreover, in these latter cases, outflow started shortly before bottom center (180 deg crank angle). In this region Fig. 4 shows that A_v/A_p for outflow was appreciably smaller in the B&S engine than in the CFR engine. Under these conditions the inlet valve of the B&S engine would have the action of a "check valve," and the outflow per unit piston area from the B&S engine would therefore be less than from the CFR engine. Consequently, the difference in the volumetric efficiency of the two engines would be smaller than if the effect of heat transfer alone was important, as was the case with cam A. Also, the relative volumetric efficiency, e_v/e_{v0}, with long inlet pipes would be less for the B&S engine than for the CFR engine, except at very low engine speeds, as shown in Fig. 7.

This explanation of Figs. 7 and 8 may be regarded as tentative in the absence of more conclusive experimental and theoretical data. It would be of considerable interest to determine whether under the condition of identical A_v/A_p versus θ, better agreement between e_v/e_{v0} of the two engines could be obtained.

Effect of Size With $(A_v/A_p)_{avg}$ Allowed to Vary. Fig. 10 shows the effect of $(A_v/A_p)_{avg}$ on the dynamic characteristics of the inlet systems of the B&S and CFR engines. Here it is seen that with a fixed inlet-valve timing, e_{v0} of the B&S engine was the same at the same gulp factor. This is in agreement with the results of reference (2), and also with the theoretical considerations presented earlier in connection with Equation [15]. On the gulp-factor scale, e_{v0} of both the B&S and CFR engines also had the same trend of variation, although the CFR engine continued to have higher volumetric efficiencies than the B&S engine, as noted in Fig. 8(*a*).

With long inlet pipes, Fig. 10 shows that at the same $(A_v/A_2)_{avg}$ and A_2L/A_pS, the e_v/e_{v0} of the B&S engine was consistently lower than that of the CFR engine, especially in the case of the smaller inlet-pipe diameter [$(A_v/A_2)_{avg} = 0.40$] of long lengths ($A_2L/A_pS = 0.91$ and 1.19), at the lower end of NL/c_0. This was probably due to the effect of greater fluid friction in the B&S engine (low

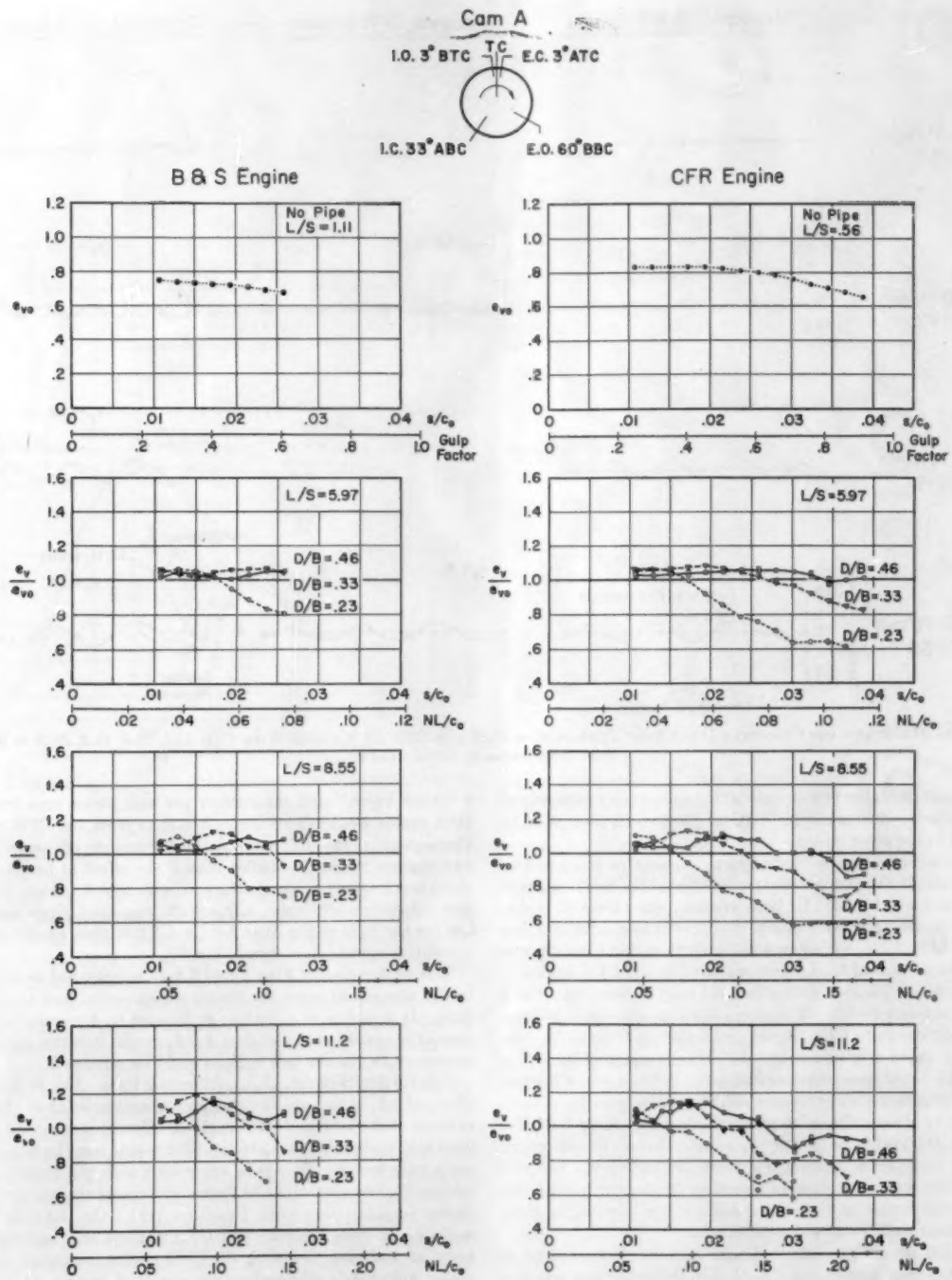


FIG. 6 VOLUMETRIC EFFICIENCIES OF B&S AND CFR (1) ENGINES OPERATING AT SAME $(A_v/A_p)_{avg}$ WITH CAM A UNDER CONDITIONS OF TABLES 1 AND 2

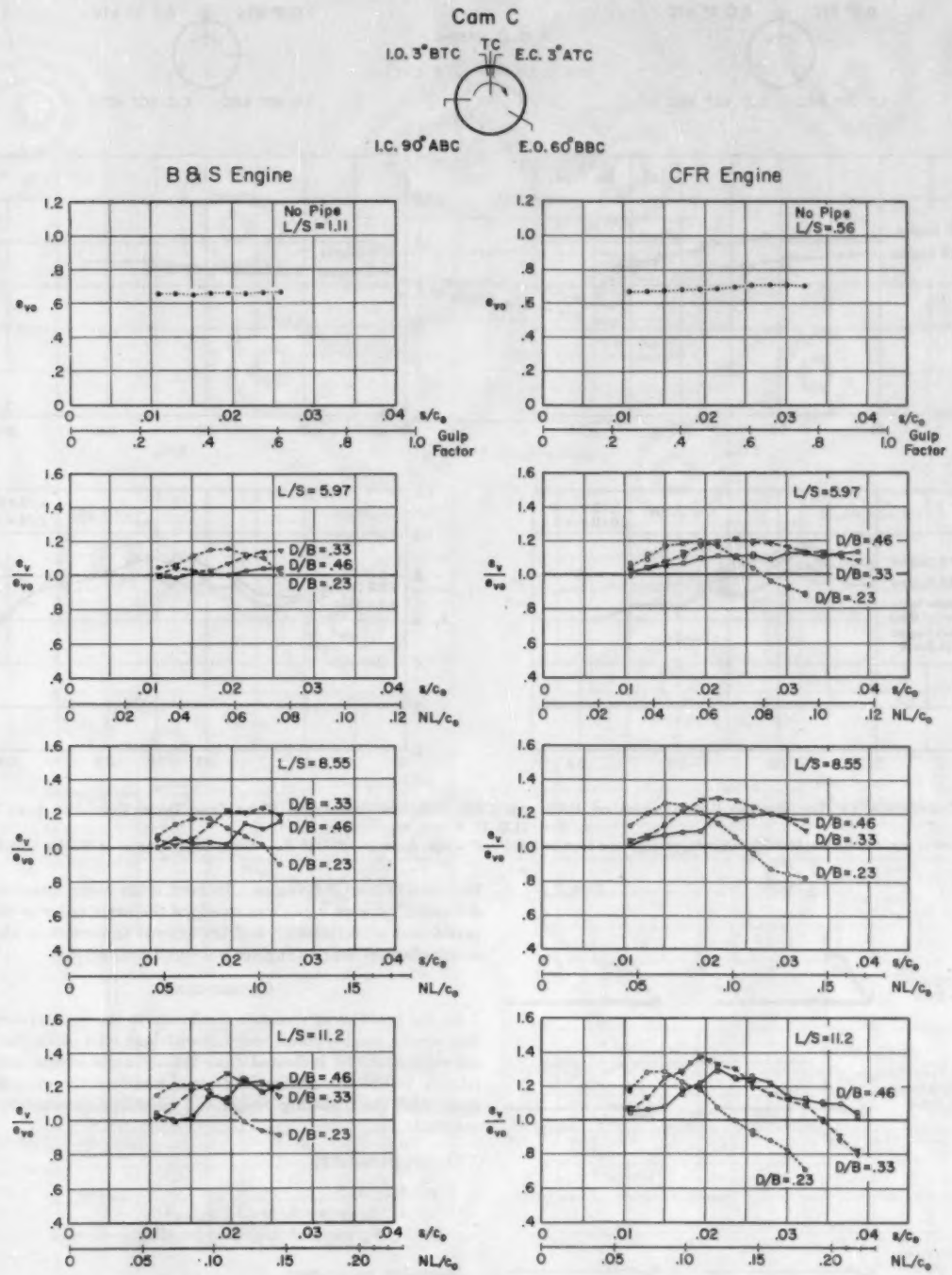


FIG. 7 VOLUMETRIC EFFICIENCIES OF B&S AND CFR (1) ENGINES OPERATING AT THE SAME $(A_v/A_p)_{avg}$ WITH CAM C UNDER CONDITIONS OF TABLES 1 AND 2

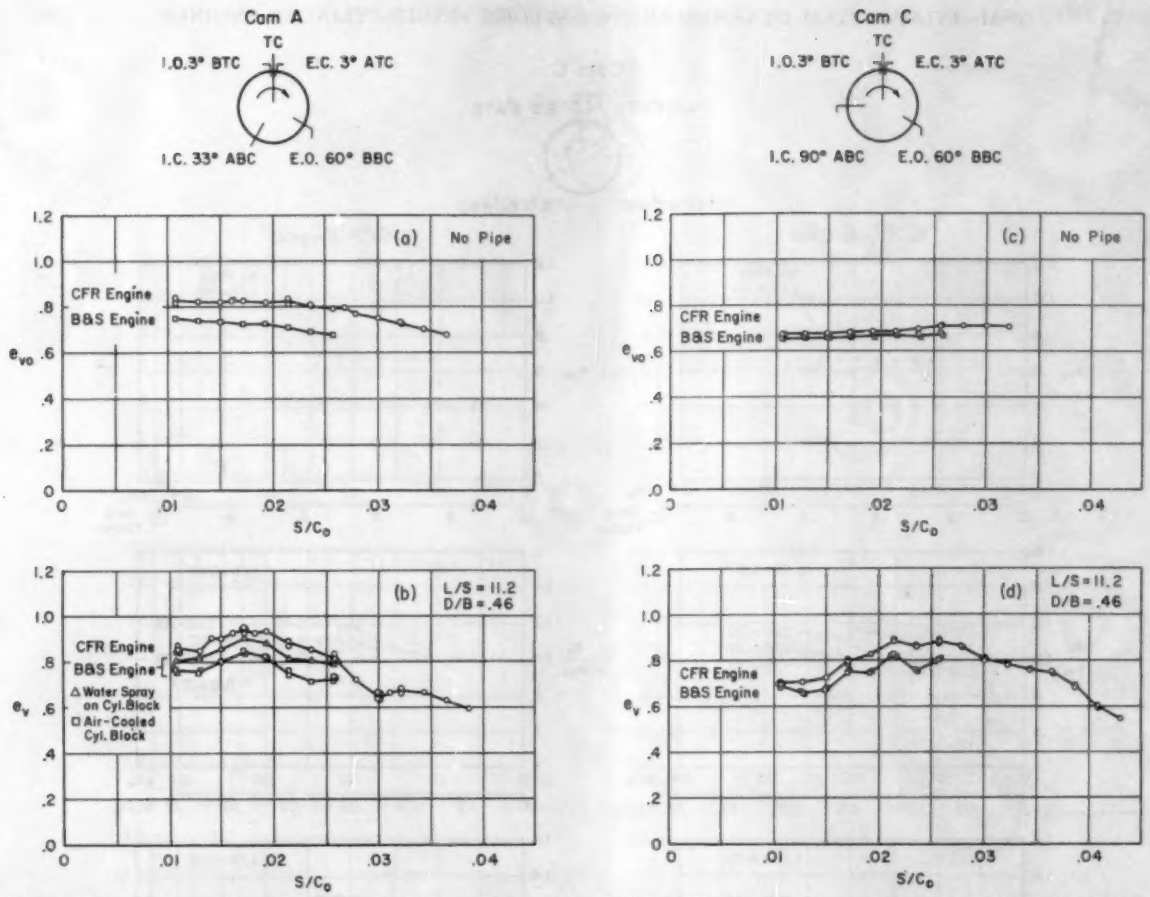
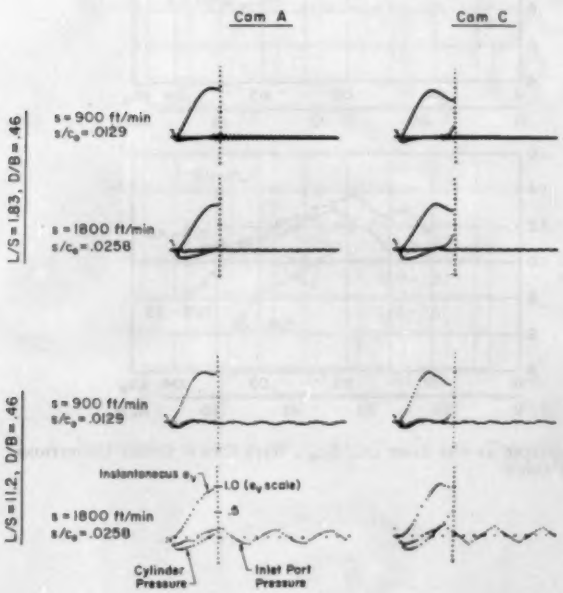


FIG. 8 COMPARISON OF VOLUMETRIC EFFICIENCIES OF B&S AND CFR ENGINES OPERATING WITH ZERO INLET PIPE, AND ALSO WITH $L/S = 11.2$, $D/B = 0.46$

(With water spray, the B&S cylinder block temperature was kept at 100 F, oil temperature at 120–140 F. Other conditions as in Tables 1 and 2.)



Re) than in the CFR engine. In most other cases, however, the difference between e_v/e_{vo} was small (of the same order as the experimental uncertainties), and the general agreement of the dynamic effect in the two engines was quite satisfactory.

CONCLUSIONS

In the problem of dynamic similarity in the inlet systems of four-stroke, single-cylinder engines with long inlet pipes, theoretical considerations indicated that the dynamic effects and the relative volumetric efficiencies (e_v/e_{vo}) would be the same in engines with the following design and operating parameters kept constant:

Design parameters

(*a*) A_2L/A_pS
(*b*) A_v/A_2 versus θ (or $[A_v/A_2]_{ave}$)
(*c*) V_c/V_d versus θ (higher-order effect)

Operating parameters

(*d*) p_a/p_0
(*e*) NL/c_0

FIG. 9 (*l*s/*l*) COMPUTED SOLUTIONS FOR CFR ENGINE (3) SHOWING EFFECT OF VALVE TIMING ON OUTFLOW FROM CYLINDER DURING INLET PROCESS

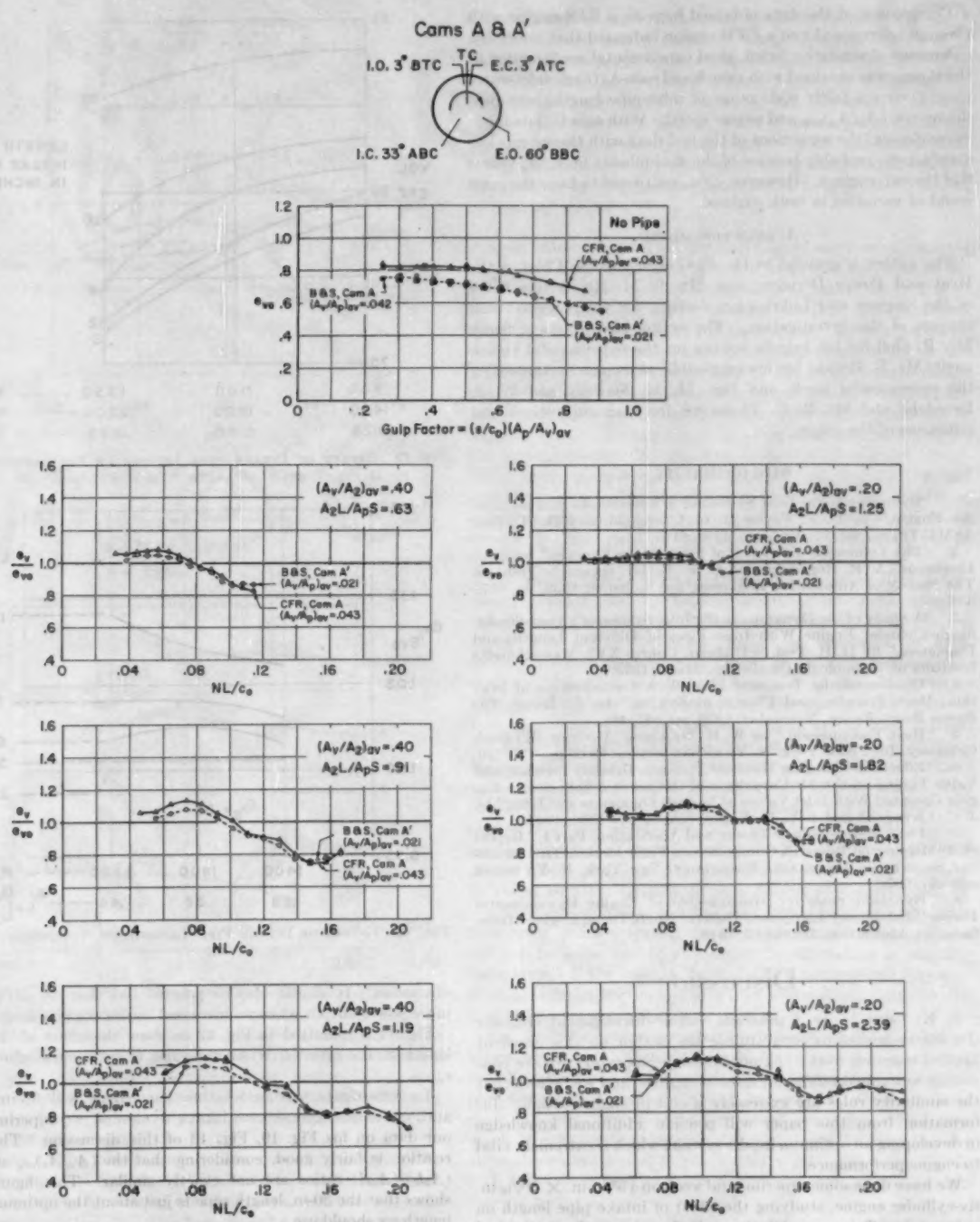


FIG. 10 VOLUMETRIC EFFICIENCIES OF B&S AND CFR ENGINES OPERATING AT DIFFERENT VALUES OF $(A_v/A_2)_{av}$
(Other conditions as in Tables 1 and 2.)

Comparison of the data obtained here on a B&S engine with those in reference (1) on a CFR engine indicated that under the conditions of similarity listed, good experimental confirmation of the theory was obtained with cam A and cam A′ (early inlet-valve closing) over a fairly wide range of inlet-pipe length, inlet-pipe diameters, $(A_v/A_p)_{avg}$, and engine speed. With cam C (late inlet-valve closing) the agreement of the test data with theory was less satisfactory, probably because of the dissimilarity in A_v/A_p versus θ of the two engines. However, e_v/e_{vo} continued to have the same trend of variation in both engines.

Acknowledgments

The author is grateful to Dr. F. G. Brickwedde, Chief of the Heat and Power Division; and Mr. P. M. Ku, Acting Chief of the Engines and Lubrication Section, for their interest and support of this investigation. The author also wishes to thank Mr. R. Goff for his help in setting up the experimental equipment, Mr. E. Bryant for his invaluable assistance in conducting the experimental work, and Drs. M. M. Slawsky and W. J. Levedahl, and Mr. R. C. Thompson for their suggestions and criticisms of the paper.

Bibliography

1 "Dynamics in the Inlet System of a Four-Stroke Single-Cylinder Engine," by C. F. Taylor, J. C. Livengood, and D. H. Tsai, ASME Trans., vol. 77, 1955, pp. 1133–1145.

2 "The Volumetric Efficiency of Four-Stroke Engines," by J. C. Livengood, A. R. Rogowski, and C. F. Taylor, Quarterly Trans. of The Society of Automotive Engineers, vol. 6, No. 4, 1952, pp. 617–634.

3 "A Study of the Dynamics in the Inlet System of a Four-Stroke, Single-Cylinder Engine With Inlet Pipes of Different Lengths and Diameters," by D. H. Tsai, ScD thesis, Course XVI, Massachusetts Institute of Technology, Cambridge, Mass., 1952.

4 "Unidimensional Transient Flow With Consideration of Friction, Heat Transfer, and Change of Section," by E. Jenny, *The Brown Boveri Review*, November, 1950, pp. 447–461.

5 "Heat Transmission," by W. H. McAdams, McGraw-Hill Book Company, Inc., New York, N. Y., second edition, 1942, pp. 117–120.

6 "Effects of Changing Manifold Pressure, Exhaust Pressure, and Valve Timing on the Air Capacity and Output of a Four-Stroke Engine Operated With Inlet Valves of Various Diameters and Lifts," by J. C. Livengood and J. V. D. Eppes, NACA TN 1366, 1947.

7 "Fluid Meters—Their Theory and Application, Part I," Report of ASME Special Research Committee on Fluid Meters, The American Society of Mechanial Engineers, New York, N. Y., fourth edition, 1937.

8 "Hydraulic Scale for Measurement of Engine Dynamometer Forces," Laboratory Equipment Bulletin No. 6, Diesel Engine Manufacturers Association, January 3, 1949.

Discussion

S. K. Chen.[7] As a practical engine-development engineer the writer wishes to congratulate the author on this excellent applied research work. At our research laboratory we are constantly working on different sizes of engines, and we believe that the similarity rules are extremely useful in engine design. Information from this paper will provide additional knowledge in developing an optimum intake system, which is extremely vital to engine performance.

We have done some experimental work on a $3^{9}/_{16}$-in. × $3^{11}/_{16}$-in. six-cylinder engine, studying the effect of intake pipe length on volumetric efficiency. All the cylinders have the individual intake and exhaust system, and therefore the intake dynamics will be similar to that of a single-cylinder engine. We varied the length of the additional intake pipe and measured the volumetric efficiency, and the results are shown in Fig. 11 of this discussion. It should also be pointed out that we did obtain more power output when we increased the intake pipe length.

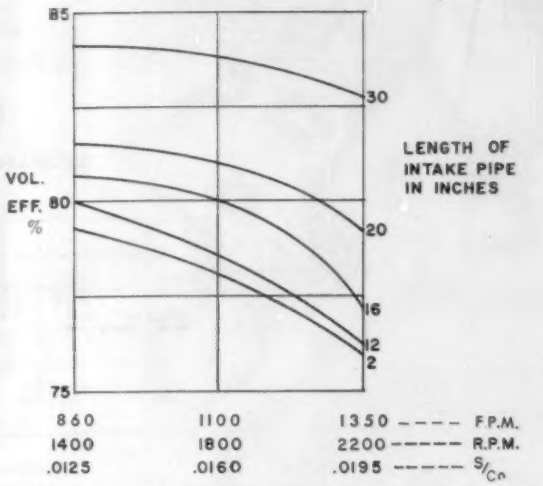


Fig. 11 Effect of Intake Pipe Length on Volumetric Efficiency

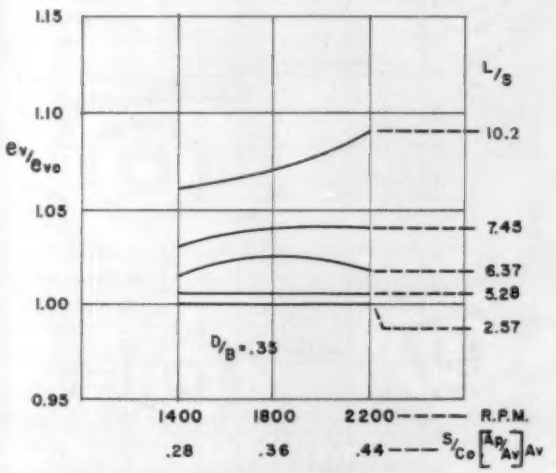


Fig. 12 Effect of Intake Pipe Length on Volumetric Efficiency

Fig. 11 is replotted in Fig. 12 to show the effect of the pipe length-stroke ratio (L/S) on the ratio of volumetric efficiencies e_v/e_{vo}.

To investigate further whether our engine will fit into the author's similarity rules on intake dynamics, we superimposed our data on his Fig. 10, Fig. 13 of this discussion. The correlation is fairly good, considering that the $(A_v/A_2)_{avg}$ and the $(A_2L)/(A_pS)$ ratios are not exactly similar. This figure also shows that the 30-in. length pipe is just about the optimum pipe length we should use.

Regarding the conditions of operating parameters, we feel that the operating parameters also should include air-fuel ratio as an important factor. The value of volumetric efficiency would not be useful unless the air-fuel ratio is specified. However, the error would be of secondary importance if the ratio of volumetric efficiencies (e_v/e_{vo}) is used to compare different engines, and if the ratio is obtained at the same air-fuel ratio.

[7] Project Engineer, Industrial Power Division, International Harvester Company, Melrose Park, Ill. Assoc. Mem. ASME.

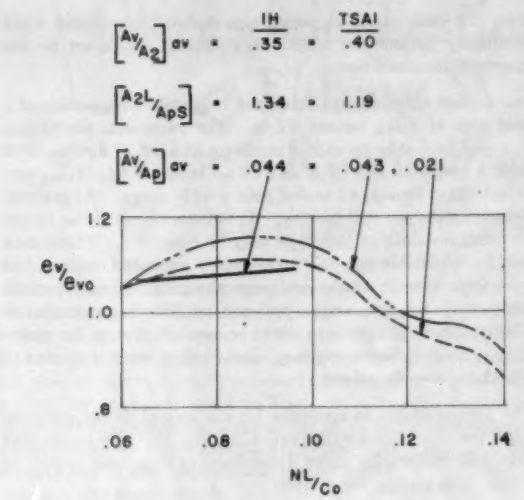


FIG. 13 COMPARISON OF I. H. DATA WITH TSAI'S DATA

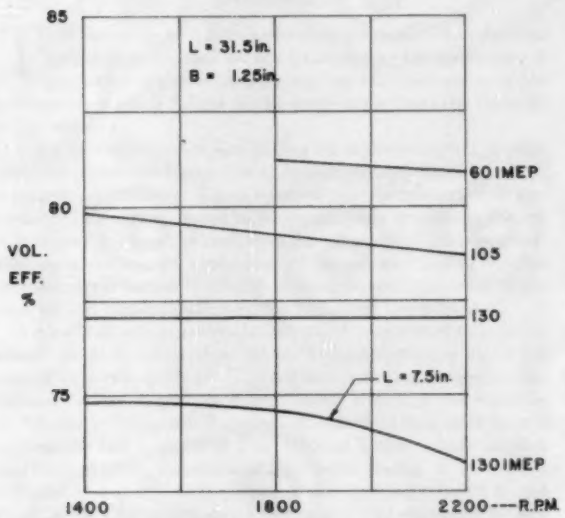


FIG. 14 EFFECT OF OUTPUT ON VOLUMETRIC EFFICIENCY

When the engine is operating at low air-fuel ratio (or higher imep), the engine is hotter and generally the whole intake system (valve, port, and intake pipe) becomes hotter. The heat transfer becomes more severe and the volumetric efficiency measured is usually lower. Fig. 14, herewith, shows the effect of imep output on volumetric efficiency.

Speaking of the intake similarity as a whole, we agree with the author that the "gulp factor" is a valid and useful parameter. In our laboratory, for simplification purposes, we define a specific flow capacity as $Q/C^{2/3}$. The flow capacity Q is defined as the total flow, experimentally determined by a specific steady flow bench test and by integration over the entire valve-lift diagram. The effect of size is taken into consideration by the $^2/_3$ power of the displacement C. When this specific flow capacity is multiplied by the operating piston speed we have a factor which is quite similar to the gulp factor.

In conclusion, it should be further emphasized that the theoretical as well as the experimental results included in the paper clearly indicate how a correctly designed intake system will yield more horsepower without a mechanical supercharger. The study of intake and exhaust systems is just as important for turbocharged engines. For example, the "blowdown" energy can be more correctly utilized if the pressure pulsation in the exhaust system is clearly understood. With this problem in mind, has there been any attempt to study exhaust dynamics by this method; and furthermore, is it feasible to apply his equations on a ported engine?

A. R. ROGOWSKI.[8] This fine paper could serve as an example to research men who apply powerful analytical methods to the solution of complex engine problems, and leave the results of their efforts in a form similar to Equation [5]. This is, to say the least, discouraging to the engine designer. In this paper, on the other hand, the author has made things much easier for the designer by interpreting his results in the form of three design parameters and two operating parameters which, if kept at the same values, will insure equal dynamic effects in new designs of different sizes. He has gone even farther by discussing and showing experimentally the separate influences of various design ratios on the magnitude of the dynamic effects. This should give the engine builder some feel for the relative importance of certain design changes and make it possible for him to predict to some extent the effect of various departures from similitude in the inlet system.

The author begins quite cautiously by assuming that A_v/A_p should perhaps also be held constant as an implication of the constant parameter A_v/A_2 versus θ. As pointed out in the paper, this added limitation makes it necessary to keep the engines and inlet systems almost completely geometrically similar. This turns out to have been a good idea, however, as Figs. 6 and 7 show quite remarkably the separate influences of L/S, D/B, and NL/c_0 on dynamic effects in the inlet system. While the magnitude and character of the influences of these various quantities are widely different, each one is seen to have an almost identical effect in engines of different sizes. The correlation is extremely good, considering the fact that the engines were not exactly similar. It is hoped that the author may be able in the future to extend the B & S curves to higher engine speeds and finally to show all the test runs, perhaps as curves of constant (A_2L/A_pS) on a master plot against NL/c_0.

A further simplification which also should be of great value is then made by investigating the assumption that dynamic effects will not be influenced to any extent by differences in effective valve area. If the same value of A_v/A_p is not required for equal dynamic effects, the engines and inlet systems can be quite dissimilar and yet be amenable to prediction.

Fig. 10 of the paper shows that the influence of A_v/A_p is indeed almost negligible. With the A′ cam, it is stated that A_2 was cut in half to keep A_v/A_2 the same. No mention of this series of pipe diameters is made in Table 1, however. If there are sufficient data now, or will be in the future, the writer would very much like to see Fig. 10 extended to include, for example, the alternate combination of $A_v/A_2 = 0.20$ with (A_2L/A_pS) of 0.63, 0.91, and 1.19 so that the separate effects of these two parameters can be better assessed.

The author is to be commended for showing how a complicated relationship may be put in a form suitable for experimental verification of the theory or evaluation of the relative importance of the controlling parameters, all with a minimum of actual engine tests.

---

[8] Assoc. Professor, Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Mass.

Author's Closure

The author is indebted to Dr. Chen and Professor Rogowski for their comments. Dr. Chen properly called attention to the effect of fuel-air ratio on volumetric efficiency. Besides the fuel-air ratio, one would expect the volumetric efficiency also to be affected by such other parameters as the thermodynamic properties of the fresh charge in the inlet system, the timing of ignition (or fuel injection), the timing and the flow capacity of the exhaust valve, etc. These parameters were not studied in the present investigation, but their influence on the *relative* volumetric efficiency e_v/e_{v0} probably would not be large, if both e_v and e_{v0} are obtained under the same conditions.

Regarding the study of exhaust dynamics by means of the theory of one-dimensional unsteady flow, there has been a continuing interest in the subject, and in the past few years, a number of papers on various aspects of the problem have been published. The author would like to note especially the recent research[9] of Wallace and Nassif on the air flow in a two-stroke engine. In their excellent paper, the authors formulated a set of similarity parameters which were closely analogous to the parameters discussed here.

The author subscribes to Professor Rogowski's suggestion of a master plot of e_v/e_{v0} versus NL/c_0. For maximum usefulness, such a plot probably should be made up of a set of figures, with curves of constant A_2L/A_pS as well as those of $(A_v/A_2)_{avg}$ and constant valve timing, all varied over a wide range. At present, systematic data are still lacking. It is hoped that in the future more complete data of this type may be compiled. These data should be applicable not only to normally aspirated engines, but also to engines with mechanical supercharging. Under certain conditions, if the interaction between neighboring cylinders is not large, these data perhaps could be applied even to the manifolding of multicylinder engines. But further work is needed to verify these considerations.

The author wishes to apologize for the lack of clarity concerning the test conditions with cam A′ in Fig. 10. He hopes that his revised paragraphs under Procedure, as they now appear in the text, help clarify the matter.

[9] "Airflow in a Naturally Aspirated Two-Stroke Engine," by F. J. Wallace and M. H. Nassif, *Proceedings of The Institution of Mechanical Engineers*, vol. 168, no. 18, 1954, pp. 515–544.

Ultrasonic Testing of Small-Diameter Tubing With Automatic Recording Equipment

By W. L. FLEISCHMANN[1] AND H. A. F. ROCHA,[2] SCHENECTADY, N. Y.

This paper presents the progress made in the ultrasonic test equipment which now appears to make automatic testing of small-diameter tubing a distinct possibility. It then presents some cross sections, showing the type and size of discontinuities detected with this equipment in Type 347 stainless-steel tubing. Finally, the summary points out the inherent difficulties in judging the severity of the defect by the amplitude indications alone and the importance of the experience factor which is required to interpret the test results.

Introduction

THE selection of a nondestructive acceptance test for materials is predicated on the likelihood of the occurrence of particular types of defects and on the seriousness of the consequences which failure of the parts made from the material might cause.

Until only a few years ago, piping for high-pressure high-temperature steam received a visual inspection and, when specified, a magnetic-particle, or a penetrant-oil test for nonmagnetic materials. This test appeared to be an adequate procedure, though it had the inherent drawback that, for all practical purposes, outside surfaces alone were inspected. The inside surface of the pipe was examined by an inspector looking through the bore of the pipe while a helper illuminated the bore from the other end.

A pipe failure in a power station, which was traced to a defect present prior to fabrication, led to the introduction of ultrasonic testing of power piping.[3] The ultrasonic-testing procedure developed for the inspection of such heavy wall pipe was reported by Moriarty[4] before this Society. The method described by him applied the test to pipes of 2 in. OD and larger. Tash[5] applied this method to the inspection of 2-in. boiler tubing.

Recently, a model heat exchanger was thermally cycled to test its suitability for liquid-metal service and was examined by sectioning the tube sheets. Inspection revealed a straight-line red-dye indication 1/2 in. long in the roller expanded section of one tube (material Type 347 stainless steel—tube dimensions 0.750 in. OD, 0.049 in. wall thickness). The metallographic examination of the cross section showed a wedgelike opening of the inner surface, Fig. 1. The sharp botton of this original discontinuity was a sufficiently high stress concentration from which the transcrystalline fatigue crack could start. Whether this fatigue crack was caused by the roller-expansion process, or propagated from the original discontinuity during the thermal-shock test is not known.

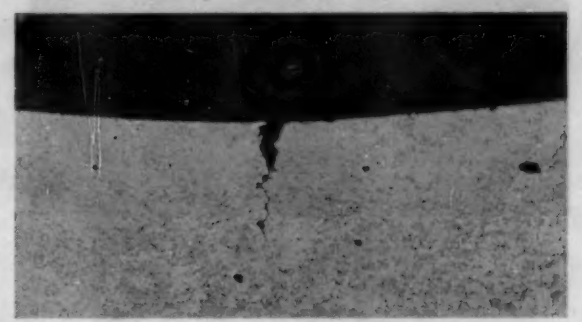


FIG. 1 FATIGUE FAILURE STARTING FROM AN ORIGINAL DEFECT AT INNER SURFACE OF PIPE

The discovery of this defect made it advisable to apply a more stringent inspection procedure to small-diameter tubing destined for atomic-power applications than was required by the appropriate ASTM specifications. It was believed that the test with the best chances of success would utilize ultrasonic shear waves.[4] Yet, the extension of a rapid ultrasonic test method to tubular-mill products depended upon the development of suitable instrumentation.

Instrumentation

The instrumentation problem was twofold. The following were needed:

(*a*) A special search unit for introducing ultrasonic pulses into the material properly.

(*b*) An automatic recording system that would reproduce faithfully the amplitude of the received echoes.

These two components will be discussed because of their general usefulness to the industry.

Delayed Shear Wave Search Unit

As recognized previously, the most effective method of testing pipe and tubing for longitudinal defects makes use of pulsed ultrasonic shear waves. These pulses are introduced at an angle by a search unit, and travel around the object causing an echo whenever they meet a discontinuity. The echoes are received by the search unit and appear as "pips" or "indications" on a cathode-ray tube. In the absence of a discontinuity, the pulses travel all the way around the object and produce a constant pip called the "end return."

---

[1] Welding Engineer, Materials Engineering, Knolls Atomic Power Laboratory, operated for the United States Atomic Energy Commission by the General Electric Company. Mem. ASME.

[2] Materials and Processes Laboratory, Large Steam Turbine-Generator Department, General Electric Company.

[3] "Report on Split Steam Lead at New Orleans," by A. W. Rankin and W. A. Reich, ASME Paper No. 50—A-76.

[4] "Ultrasonic Flaw Detection by Means of Shear Waves," by C. D. Moriarty, Trans. ASME, vol. 73, 1951, pp. 225–235.

[5] "Field Inspection of Boiler Tubes With Ultrasonic Reflectoscope," by J. A. Tash, Trans. ASME, vol. 74, 1952, pp. 201–206.

Contributed by the Metals Engineering Division and presented at the Diamond Jubilee Spring Meeting, Baltimore, Md., April 18–22, 1955, of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS.

NOTE: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, January 17, 1955. Paper No. 55—S-23.

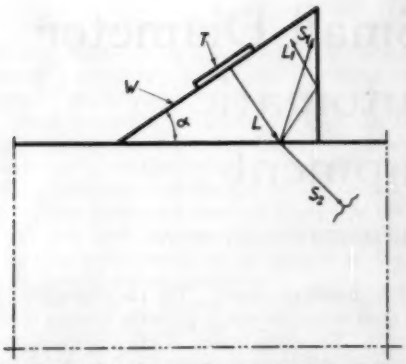


FIG. 2 SCHEMATIC PICTURE OF STANDARD SHEAR WAVE SEARCH UNIT

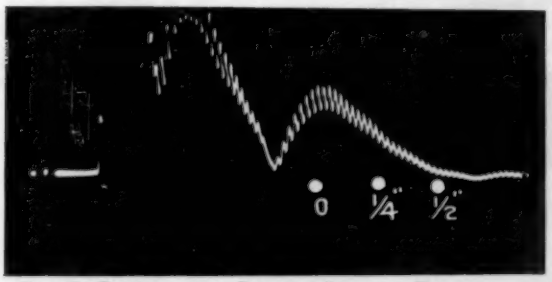


FIG. 3 REFLECTOGRAM SHOWING INTERNAL REFLECTIONS IN STANDARD SEARCH UNIT

The standard shear wave search unit, Fig. 2, consists of an electromechanical transducer T, mounted on a plastic wedge W. The longitudinal wave L generated by the transducer is refracted and reflected at the wedge-object interface. The refracted beam, in general, has two components—longitudinal and shear. However, the angle of incidence is made such that the longitudinal component is totally reflected. As a result, only the shear wave S_2 is transmitted into the object.

The reflected beam, with its two components, the longitudinal L_1 and the shear S_1, bounces from wall to wall in the wedge W, producing a series of reverberations. These are received by the transducer T and shown on the cathode-ray tube as an extension of the initial pulse.

When testing a large object, in which the defects are expected to be far enough away from the search unit, this phenomenon is not objectionable. However, this is not the case when the search unit must, of necessity, be close to the defects.

In the case of small-diameter tubing, these reverberations extend up to the end return, as shown in the reflectogram of Fig. 3. Under this condition, small defects would pass undetected.

The use of a longer sweep (including multiple end returns) was not considered practical because dispersion and scattering of the echoes increases with the distance traveled, thus making the detection of small defects unreliable.

The problem finally was solved by the use of a "delayed shear wave search unit," previously developed for the purpose of detecting defects lying near the testing surface.

In this search unit, Fig. 4, the transducer T is mounted on a "delaying wedge" D which delays the entrance of each pulse into the object by a time interval t_d which is larger than half the duration of the pulse itself t_p

$$t_d > \frac{t_p}{2}$$

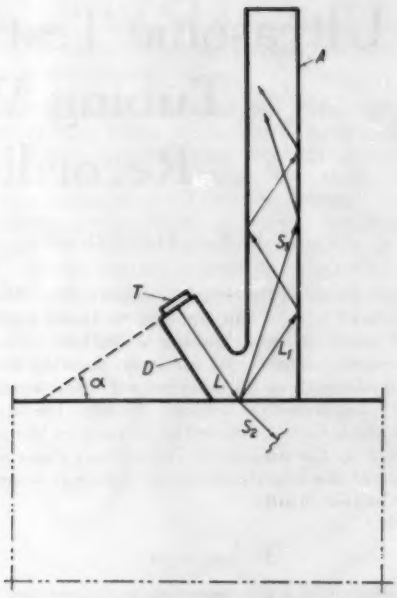


FIG. 4 SCHEMATIC PICTURE OF DELAYED SHEAR WAVE SEARCH UNIT

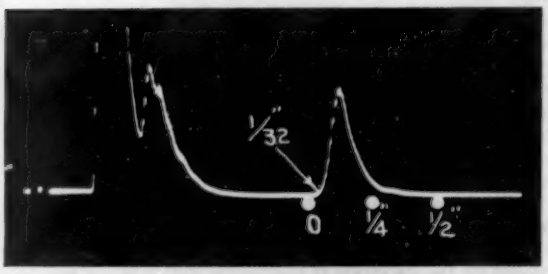


FIG. 5 REFLECTOGRAM OF DEFECT LOCATED 1/32 IN. FROM DELAYED SHEAR WAVE SEARCH UNIT

Thus, if a defect next to the surface produces an echo that is reflected toward the transducer, it will arrive after t_p at which time the transducer has ceased to vibrate and the receiver is capable of amplifying it properly.

To eliminate the annoying reverberations present in the conventional search unit (which, as said before, obscure the sweep for a considerably long time after the transmitted pulse ceases), an "absorbing wedge" A is incorporated in the delayed shear wave search unit. In this absorbing wedge both components L_1 and S_1 of the reflected beam are forced to bounce along the walls, losing energy by absorption as they travel along. At high enough frequencies (in this case 5 Mc/sec) this absorption is such that no reverberation is perceptible at normal testing sensitivities.

The clean sweep produced by this type of search unit is evidenced in the reflectogram of Fig. 5 in which an artificial defect (hole no. 60 drill, 1/4 in. deep) is detected 1/32 in. below the surface.

By fitting the contact surface of this search unit to the curvature of the tube to be tested, a clean sweep covering 180 deg is obtained.

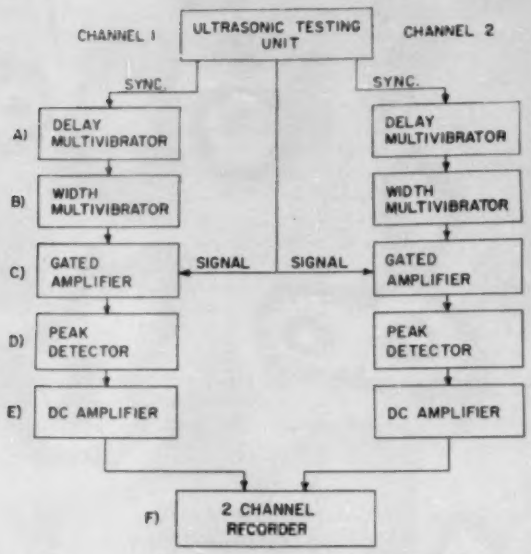


FIG. 6 BLOCK DIAGRAM OF ULTRASONIC RECORDING UNIT

AUTOMATIC RECORDING SYSTEM

A proper recording of this type of ultrasonic test must indicate:

1 The magnitude of any indication showing in the time interval between the end of the initial pulse and the beginning of the end return.

2 The amplitude of the end return, to make sure that the testing sensitivity remains constant during each test.

These specifications are met by the automatic recording unit shown in the block diagram of Fig. 6. This recording unit comprises two independent channels which are identical, with the exception that they are "active" at different times.

Channel 1 is "on" during the period between the end of the initial pulse and the start of the end return.

Channel 2 is on during the period in which the end return is being received.

The timing of the electronic "gates" that turn each amplifier on and off at the proper moment is controlled by two multivibrators for each channel. The "delay multivibrator" A sets the time at which the gate is turned on. The "width multivibrator" B sets the time at which the gate is turned off.

Any echoes arriving at the gated amplifier C, during the time the gate is on, are amplified; the ones arriving during the off period are rejected. The output of the gated amplifier is fed to a peak-detector D, which measures the maximum amplitude of each pulse; after further amplification through a d-c amplifier E, the results are fed to a conventional two-channel electromagnetic recorder F.

Sensitivity of channel 1 is set by making the indication from a standard artificial defect (3 per cent notch) equal to a predetermined deflection of the recorder needle.

The amplitude of the end return (channel 2) gives a constant measure of the test sensitivity. If, for instance, the coupling between the search unit and the tube decreases, the amount of energy transmitted into the tube is reduced and a correspondingly lower amplitude of the end return is recorded.

Besides the advantage of printing a permanent test record of each item, this type of instrumentation offers further approaches

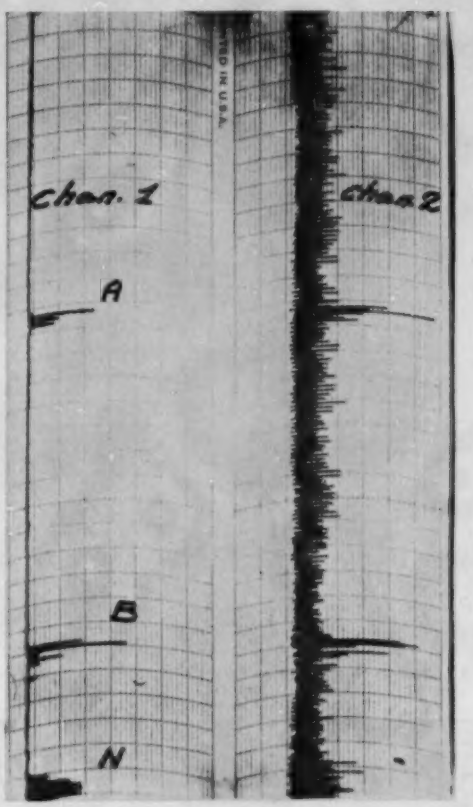


FIG. 7 REFLECTOGRAM RECORDED ON A TWO-CHANNEL ELECTROMAGNETIC RECORDER

to automation. If an indication is present with an amplitude exceeding the one produced by the standard artificial defect, a trigger circuit can be made to operate an audible or visible alarm system and set off a spray gun to paint the location of the defect. If loss of coupling causes the end return to drop below a minimum value, another alarm can be sounded and the equipment brought to a stop.

TEST RESULTS

In Fig. 7 is shown a typical test record of a defective tube. In the left-hand side (channel 1) the three indications shown were produced by the following:

1 A 3 per cent sensitivity notch at point N.
2 Two discontinuities at points A and B.

In channel 2 the amplitude of the end return is recorded, showing that a reasonably constant contact was maintained during the test. Naturally, at points A, B, and N, at the moment the indications are traveling across the end return, its amplitude increases above the normal level. This occurs when the defect is about 180 deg away from the transducer.

This tube (Type 347 stainless steel, 1 in. OD, 0.085 in. wall thickness) was cross-sectioned at points A and B. Metallographic examination revealed in each case a seam on the inside of the tube (see insert in Fig. 8). The depth of these seams is shown in the cross section of Fig. 8. They measure, respectively, 8 and 9 per cent of the wall thickness.

The length of the discontinuity is not a critical factor. Defects

much smaller in length than the transducer, such as pinholes, have been detected easily.

The sensitivity of this test is well illustrated by the defects shown in Figs. 9 and 10.

Fig. 9 shows a spot from which a short, shallow sliver was torn.

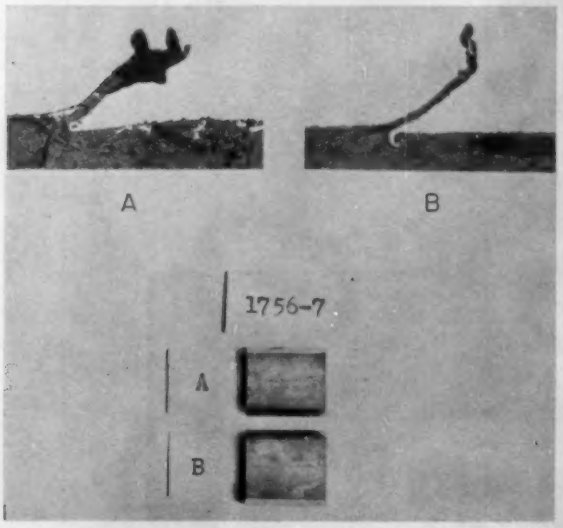


FIG. 8 DISCONTINUITIES FOUND BY ULTRASONIC TEST IN 1-IN-OD 0.085-IN-WALL TYPE 347 STAINLESS-STEEL TUBING

(a)

In the photomicrograph of Fig. 10 a small lamination is shown. This lamination had a circumferential length of only 0.02 in. and produced an ultrasonic indication which, although small, was definitely detectable.

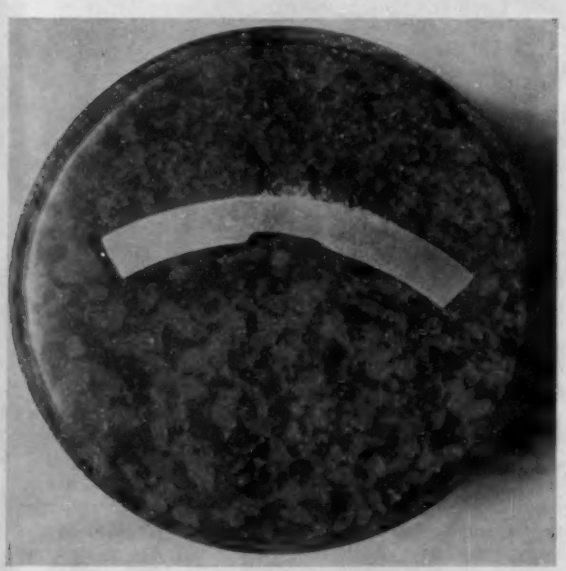
In all these tests, a 3 per cent milled notch was used as the standard indication. However, the amplitude of each indication is a function not only of the size of the defect, but also of its shape and angular position in relation to the scanning beam. Thus a comparatively large defect may produce a small indication. To minimize this condition, scanning in two directions is recommended. An ultrasonic testing unit, if used with two automatically switched search units, can perform the double scanning operation in the time normally required for single scanning.

The evaluation of the test results depends upon two factors: (1) The alloy and the method of manufacture influence the types of discontinuities which are usually found in the inspected part. This experience can only be accumulated by metallurgical examination of the areas in which ultrasonic indications are found. (2) The type of service into which the products will go will dictate the severity of the inspection.

SUMMARY

The review of the principles involved in the design of the delayed shear-wave search unit shows the solution of one of the basic

FIG. 10 0.02-IN-LONG LAMINATION ALSO FOUND BY ULTRASONIC TEST



(b)

FIG. 9 1-IN-LONG DISCONTINUITY DETECTED BY ULTRASONIC TEST ON INSIDE OF 1-IN-OD TUBE

problems encountered in the extension of the ultrasonic testing to small objects.

The introduction of automatic recording is another step to increase the usefulness of ultrasonic testing. It makes the operation independent of the human variables always present with visual observation of the screen.

In this discussion the application of ultrasonics to product inspection has been stressed. However, it is believed that the greatest usefulness of the equipment described will be as a quality-control tool at the producer's plant tracing defects to the manufacturing step at which they first occur.

Acknowledgments

The authors take pleasure in acknowledging the contributions of Messrs. R. E. Warnow and M. E. Auger, Materials and Processes Laboratory, Large Steam Turbine-Generator Department, General Electric Company, in the construction and testing of the ultrasonic equipment. Special thanks are due to Mr. R. D. Brooks of the Atomic Power Equipment Engineering Section, General Electric Company, who was instrumental in putting the equipment into practice, and also to Messrs. A. L. Williamson and E. Van Zylstra who followed the application.

Addendum

After the manuscript was submitted, the important question was raised[6] whether a pipe or tube with eccentric walls can produce indications similar in appearance to those normally produced by discontinuities.

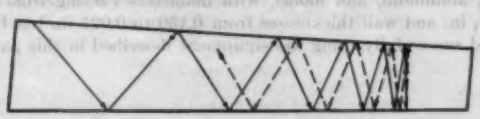


Fig. 11 Formation of a False Echo

The curious though logical answer is yes. Fig. 11 shows a constant slope wedge into which a wave has been introduced at 45 deg. The angle of incidence decreases by a constant value for each reflection; after a few reflections the wave reverses itself and starts traveling back toward the initial point. Thus an echolike indication is formed which is displayed on the face of the cathode-ray tube. However, this false echo is formed only when the wave travels in the direction of decreasing thickness.

In the case of a tube with eccentric walls, the situation is equivalent to two wedges with slopes of opposite signs. When the wave is traveling in the part with decreasing thickness, a false echo may be formed which can easily be mistaken for a defect indication.

However, as can be seen from Fig. 11, it takes several reflections before the direction of the wave is reversed and the false echo is returned. Hence no indication will be present in the initial portion of the sweep.

With the conventional search unit this cannot be observed because the initial pulse and wedge reverberations obscure that initial part of the sweep. However, with the delayed shear wave search unit, the zero point of the sweep (corresponding to the point of incidence of the shear-wave pulse) is perfectly clean and, as a result, that phenomenon can be observed.

While the eccentric tube rotates, the false echoes can be seen traveling across the screen, but contrary to true defect indications they die out before they reach the zero point of the sweep. In this way, a clear distinction between true and false indications is accomplished. Moreover, for automatic recording, gate No. 1 can be set next to the zero point of the sweep and, in that way, made immune to false echoes.

[6] Mr. Frank Lambert, Oak Ridge National Laboratory, private Communication.

Discussion

J. J. B. Rutherford.[7] Inspection of steel by ultrasonic methods has found widespread prevalence in the past few years, but it has generally been restricted to the inspection of large masses of metal containing machined surfaces. The authors have devised an ingenious method of damping out overtones and thereby making the technique applicable to mill-produced parts.

Since this method of inspection has been developed for application to a specific case of close-tolerance tubing, it would be interesting to know the modifications necessary for variations of this product. For example, what modifications of transducer and wedge are necessary with change of tube diameter—change of wall thickness—change from Type 347 material?

L. N. Wall.[8] The authors are to be congratulated on their clear, concise presentation of this paper free of lengthy description of electronic circuits, which, after all, is not essential to a proper understanding of the operation of their equipment. Their equipment fills a need long felt in industry for a reliable, fast method of nondestructive testing of small-diameter pipe and tubing for internal defects and also provides a convenient permanent record of the test results.

In co-operation with the Knolls Atomic Power Laboratory, we have tested several thousand feet of tubing using the equipment described in this paper. The tubing tested was $1^1/_4$ and $1^1/_2$ in. OD $\times$ 0.065 and 0.095 in. minimum wall thickness of both plain carbon and AISI Type 347 steel. A semiautomatic device for holding the searching unit in contact with the tubing and for revolving and advancing the tubing at a predetermined rate past the searching unit was used in all our tests. The fluid couplant was also automatically fed under the searching unit.

As a standard defect, we used a milled longitudinal notch on the inside surface of a piece of tubing of the same material and size as the tubing to be tested. Although in theory, the angle the side walls of the notch make with the incident sound beam should affect the magnitude of the back reflection from the notch, we found no measurable difference in the strength of the back reflection from a vee and a square notch. For that reason we standardized on a straight side-wall notch.

Using as a basis for rejection any defect which gave an indication larger than that from a milled notch whose depth was equal to 3 per cent of the wall thickness of the tubing, we found as high as 30 per cent of some lots of tubing rejectable. The majority of these defects were found to be mandrel marks at the inside surface. The remainder were laps or seams also at the inside surface. In the early stages of this test program, all tubes which gave an indication of a defect 3 per cent of the wall thickness were sectioned at the apparently defective area. In every case, actual defects exceeding 3 per cent were found by visual examination in the sections. However, it was found that the defect indications were not necessarily proportional to the actual size of the defect. This is to be expected since the orientation and surface condition of the defect, as well as size, will affect the ultrasonic indications.

It is hoped that as further experience is obtained in industry with the ultrasonic testing it will be possible to correlate more closely the signal indication with the size and shape of the defect.

[7] Metallurgist, Tubular Products Division, The Babcock & Wilcox Company, Beaver Falls, Pa. Mem. ASME.

[8] Metallurgical Division, Combustion Engineering, Inc., Chattanooga, Tenn.

However, for the present where it is imperative that internal defects in tubular materials must be detected, this equipment provides the only method of detection for long lengths of magnetic and nonmagnetic tubular materials and should find a wide use.

Authors' Closure

We thank both discussers for their pertinent remarks. We feel encouraged by the progress which has been made and to which Babcock & Wilcox and Combustion Engineering, Inc., have contributed.

In regard to Mr. Rutherford's questions concerning variations of the search unit due to changes in material and size of tubing, the following comments are in order:

(*a*) *Material.* The area of contact, as shown in Fig. 12 of this closure, must be carefully chosen in order to avoid transmission of longitudinal waves and formation of strong surface waves.

The first condition is met by making the minimum angle of incidence α_1, larger than the first critical angle α_{c1}

$$\alpha_1 > \alpha_{c1} = \sin^{-1} \frac{V_{L1}}{V_{L2}}$$

in which V_{L1} and V_{L2} are the velocities of longitudinal waves in the search-unit wedge and in the tubing.

The second condition (no strong surface waves) is met by making the maximum angle of incidence α_2, smaller than the second critical angle α_{c2}

$$\alpha_2 < \alpha_{c2} = \sin^{-1} \frac{V_{L1}}{V_{S2}}$$

in which V_{S2} is the velocity of shear waves in the tubing.

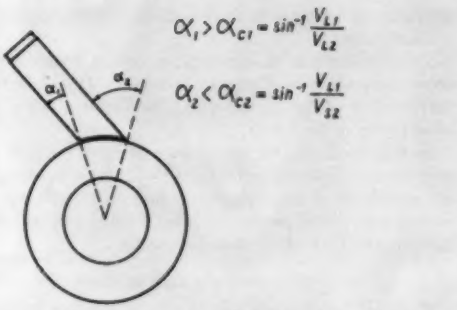


Fig. 12 Determination of Angles of Incidence of Search Unit

It is apparent from the foregoing formulas that materials with different sound velocities require different angles of incidence.

(*b*) *Outside Diameter.* A variation in outside diameter requires a corresponding variation in the curvature of the area of contact for proper transfer of energy; at the same time, in order to keep α_1 and α_2 constant, the length of the arc of contact must vary in direct proportion to the diameter.

(*c*) *Wall Thickness.* The considerations under (*a*) and (*b*) apply to the case in which the wall thickness is large in relation to the wave length. If this is not the case, Lamb waves will be formed and the optimum angle of incidence must be determined experimentally.

(*d*) *Actual Practice.* Tubing made of ferrous steel, stainless steel, aluminum, and monel, with diameters varying from $1^1/_2$ to $^1/_4$ in. and wall thicknesses from 0.150 to 0.025 in. has been tested successfully using the equipment described in this paper.

Heat Transfer to Turbulent Boundary Layers With Variable Free-Stream Velocity

By R. A. SEBAN[1] and D. L. DOUGHTY[2]

Local heat-transfer coefficients have been determined for the turbulent boundary layer of a two-dimensional subsonic air flow over the surface of a flat plate, with the air velocity variable over the plate length. The boundary layer, formed with almost constant velocity, was then subjected to a strong acceleration which produced a three to fourfold increase in the free-stream velocity, and after this maximum there was a gradual deceleration of the flow. Maximum velocities ranged from 200 to 900 fps. The use of the results, obtained for a particular kind of variation of velocity, is primarily as a check of methods of predicting heat transfer under circumstances of this type. It is shown that adequate predictions can be made from an analysis based on constant properties, and that the most important element of the analysis is an account of the thermal energy in the boundary layer at any point along the plate. Results are given also for constant velocity along the plate, both for a completely turbulent flow and for turbulent flow occurring after natural transition. The heat-transfer coefficients so obtained for the turbulent boundary layer agree best with the prediction of the von Karman analogy.

Nomenclature

The following nomenclature is used in the paper:

c = constant pressure specific heat, Btu/slug deg F
c_f = local skin-friction coefficient
h = local heat-transfer coefficient, Btu/hr sq ft deg F
H = ratio of displacement to momentum thickness
k = thermal conductivity, Btu/hr sq ft deg F/ft
L = length (1 ft in all cases dealing with the plate)
q = heat flux at the wall, Btu/hr sq ft
Q = total heat transfer to point x, Btu/hr ft width
t = temperature, deg F
t_{aw} = adiabatic wall temperature, deg F
t_0 = wall temperature, deg F
t_s = stagnation temperature, deg F
u = velocity, fps
u_1 = free-stream velocity fps
u_a = reference velocity, taken as velocity at station 0.234 ft from leading edge
x = distance from leading edge on surface of plate, ft
x_h = heated length, ft
y = distance normal to surface (mils = 10^{-3} in. used on figures)
θ = momentum thickness, ft
δ^* = displacement thickness, ft
ν = kinematic viscosity, ft²/sec
ρ = density, slugs/ft³
σ = Prandtl number

Introduction

To obtain information on the effect of variable free-stream velocity on the rate of heat transfer to turbulent boundary layers, observations have been made on a heated flat plate mounted in the test section of a wind tunnel. Longitudinal variation of the free-stream velocity was obtained by placing blisters on the tunnel walls, and completely turbulent flow existed through the agency of grit roughness elements placed near the leading edge of the plate. Without the blisters, the velocity was practically constant along the length of the plate, and results also were obtained for this condition. Fig. 1 reveals some typical velocity distributions obtained with and without the blisters, and shows also the position of the plate and the blisters in the test section of the wind tunnel.

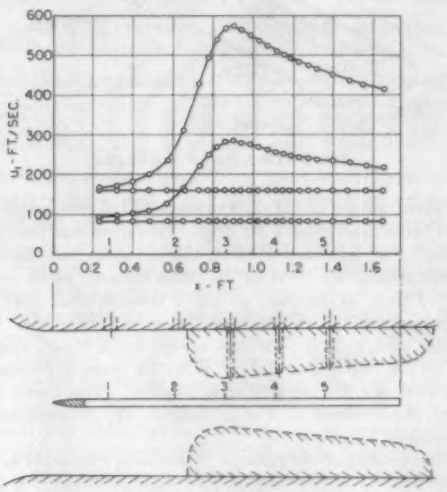


Fig. 1 Typical Free-Stream Velocity Distributions and Position of Plate and Blisters in the Test Section
(The numbers 1, 2, 3, 4, 5 indicate the stations at which boundary-layer traverses were made.)

With constant free-stream velocity, the local heat-transfer coefficient has already been quite well defined experimentally in relation to the important variables that affect it. For air flows with turbulent boundary layers, the Colburn and von Karman analogy expressions give nearly the same values, and with a local friction coefficient evaluated from a suitable correlation, the predictions of the heat-transfer coefficient given by these expressions agree fairly well with experimental results. This is true also for the results presented in this paper, which were obtained at air speeds from 90 to 460 fps.

With variable free-stream velocity, on the other hand, knowledge about the heat transfer to turbulent boundary layers is not as well defined, and both theories and experimental evidence are

[1] Associate Professor of Mechanical Engineering, University of California, Berkeley, Calif. Mem. ASME.
[2] Engineer, Shell Development Company, Emeryville, Calif. Assoc. Mem. ASME.
Contributed by the Heat Transfer Division and presented at the Diamond Jubilee Semi-Annual Meeting, Boston, Mass., June 19–23, 1955, of The American Society of Mechanical Engineers.
Note: Statements and opinions advanced in papers are to be understood as individual expressions of their authors and not those of the Society. Manuscript received at ASME Headquarters, May 13, 1955. Paper No. 55—SA-68.

at a minimum. Some of the bases of prediction of the local heat-transfer coefficient that have been recommended include the use of the flat-plate correlation on a local basis, or the prediction of the local friction coefficient by an analysis based on the integral form of the momentum equation, with the local heat-transfer coefficient obtained from the friction coefficient by an analogy expression. The experimental results for the heat-transfer coefficient as obtained for variable free-stream velocity indicate that this latter recommendation is satisfactory only if the procedure is extended to include an important correction which accounts for the thermal history of the boundary layer.

Apparatus

The heat-transfer model, a plate 6 in. wide, $^1/_2$ in. thick, and 21 in. long, was mounted centrally in the 6-in. × 9-in. test section of the wind tunnel. This tunnel provided air temperatures uniform in space and time to about $^1/_4$ deg F and a turbulence level of the order of 0.3 per cent in the speed range of 200 to 400 fps, though at lower velocities this increased due to some pulsation in the flow.

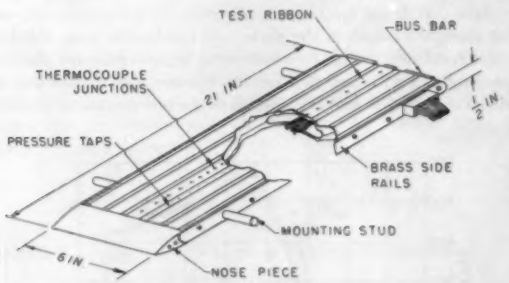


Fig. 2 Details of the Plate

The plate, shown in Fig. 2, was made of bakelite, except for a tapered brass nose piece 2 in. long. Five nichrome heating ribbons, each 1 in. wide and 0.002 in. thick, passed from the rear of the plate around the front at the point between plate and nose piece and back to the rear, at which the electrical connections were located, and the ribbons were cemented to the plate. Spaces of $^1/_{16}$ in. separated the ribbons, except for a $^1/_{10}$-in. space adjacent to the center test ribbon, in which space a pressure tap was located at each thermocouple position. The points on the velocity distributions of Fig. 1 indicate the pressure-tap and thermocouple locations.

Electric heating of the ribbons, which were connected in series, was provided by alternating current supplied through a voltage regulator.

To produce early transition to turbulence, $^3/_4$-in-wide strips of double adhesive tape were attached to each side of the nose section of the plate, completely spanning it, with the upstream edge of the tape 0.104 ft downstream from the leading edge. Number 100 crystallon grit was distributed on the face of the tape, which itself was 0.003 in. thick.

The blisters used for producing variable velocity were made of wood, and spanned the 6-in. width of the tunnel, and they were placed symmetrically to produce the same velocity distribution on both sides of the model.

Data and Results

The data obtained for all types of operation were those leading to the velocity distribution along the plate, the recovery factor, and the local heat-transfer coefficient. Recovery factors were determined from the defining equation

$$r = 1 - \frac{t_s - t_{aw}}{u_1^2/2cJ} \quad \ldots\ldots\ldots\ldots\ldots\ldots [1]$$

with the difference $(t_s - t_{aw})$ evaluated as the observed difference between stagnation and wall temperatures with the plate unheated. Thus some of the effects of radiation and internal conduction were included, but the recovery factors as obtained are estimated to be accurate to about 2 per cent when the local velocity exceeded 300 fps, and results are presented only for such conditions.

Heat-transfer coefficients were evaluated from the definition

$$h = \frac{q}{(t_0 - t_{aw})} \quad \ldots\ldots\ldots\ldots\ldots\ldots [2]$$

in which the difference $(t_0 - t_{aw})$ was the difference between the observed temperatures with heated and unheated operation, and the heat flux q was evaluated from the calculated electrical dissipation. The effects of internal conduction and radiation thus introduced errors which were generally small into the heat-transfer coefficient. Heat conduction through the plate was minimized by the symmetrical surface-temperature distribution, and spanwise conduction, at least from the center test ribbon, by the pairs of adjacent ribbons. Longitudinal conduction was significant only when a substantial variation in surface temperature occurred, and for the results presented it produced a maximum error of 3 per cent at the point of the most rapid change in temperature. This error was local, and diminished to less than 1 per cent at distances $^1/_2$ in. away from the point of maximum error. Radiation effects were small because of the combination of high heat-transfer coefficients and the low emissivity of the ribbon. All the errors, except for longitudinal conduction, taken together with the precision of the instrumentation, lead to a maximum expected error of less than 5 per cent in the local heat-transfer coefficient as evaluated from Equation [2].

All data were taken with stagnation temperatures between 80 and 100 F, and stagnation pressures near atmospheric pressure. Temperature differences between plate and stagnation temperature were of the order of 25 deg F at the point of maximum temperature difference.

Besides the heat-transfer results, some specification of the flow is presented, as obtained from impact-tube surveys of the boundary layer, made at locations indicated in Fig. 1. The impact tube had an approximately rectangular tip, 0.048 in. wide and 0.0097 in. high, with a 0.0265 × 0.0022-in. opening. Results obtained from the traverses are presented as direct evaluations from the data, without correction for such effects as the variation of velocity across the height of the tube. The kind of correction which might exist is not known certainly; a possible one would produce an outward shift of all points on the velocity distribution by 0.002 in.

Heat Transfer With Constant Free-Stream Velocity

Heat-transfer coefficients obtained with constant free-stream velocity have been adjusted for the effect of the unheated starting length produced by the brass nose piece on the model. This adjustment is uncertain because of the conduction of heat in the brass and the generation of heat in the region in which the ribbons pass between the nose piece and the bakelite plate, and also to some extent because the free-stream velocity is not constant over the entire nose piece. To apply the analytical correction, heating was taken to begin at that point which would have been reached by the ribbons if the section covered by the juncture of nose piece and bakelite plate had been extended over the surface of the model. The unheated length so defined is 0.140 ft, and the correction for a step in the heat rate at that

point, according to Klein and Tribus (1),[3] produces a reduction in the observed coefficients through multiplication by the factors in Table 1.

TABLE 1 CORRECTION FACTORS FOR EFFECT OF UNHEATED STARTING LENGTH

Station, x, ft						
	0.234	0.318	0.401	0.484	0.568	0.651
	0.734	0.776	0.818	0.848	0.900	0.925
	0.939	0.949	0.957	0.962	0.964	0.966

Fig. 3 shows the Stanton numbers, evaluated from the adjusted heat-transfer coefficients, as a function of the Reynolds

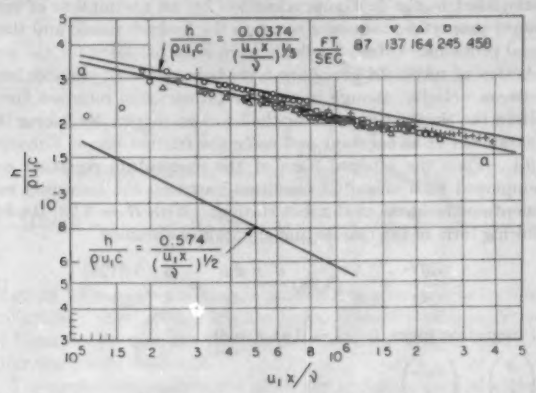


FIG. 3 STANTON NUMBERS FOR CONSTANT FREE-STREAM VELOCITY
(Curves indicated by equations are theoretical result for laminar flow and constant heat rate and the Colburn relation for turbulent flow. Curve a is the von Karman analogy. All properties contained in this representation were evaluated surface temperature and pressure.)

number, the conventional representation of heat-transfer results for this kind of operation. For comparison, there are shown the Colburn analogy

$$\frac{h}{\rho u_1 c} = \frac{c_f}{2}\frac{1}{\sigma^{2/3}} \quad [3]$$

and the von Karman analogy

$$\frac{h}{\rho u_1 c} = \frac{\sqrt{\frac{c_f}{2}}}{5\sigma + 5\log(5\sigma + 1) - 14 + \sqrt{2/c_f}} \quad [4]$$

Both are plotted for a Prandtl number of 0.70 and a local friction coefficient obtained from the relation

$$c_f = \frac{0.059}{\left(\frac{u_1 x}{\nu}\right)^{1/5}} \quad [5]$$

The theoretical result for a laminar boundary layer on a plate with constant heat flux at the surface is also shown for comparison.

Comparison of the experimental results to these predictions for turbulent flow presumes that the boundary-layer flow conforms to that for which the expression for the skin-friction coefficient applies. This was examined for one speed by means of impact-tube surveys of the boundary layer, which produced the mean speed distributions shown in Fig. 4. Such distributions can be analyzed, in the way suggested by Clauser (2), to infer the friction coefficient which would make them fit the universal semilogarithmic distribution in the region near the wall. Table 2 contains such values obtained from the velocity distributions of

[3] Numbers in parentheses refer to the Bibliography at the end of the paper.

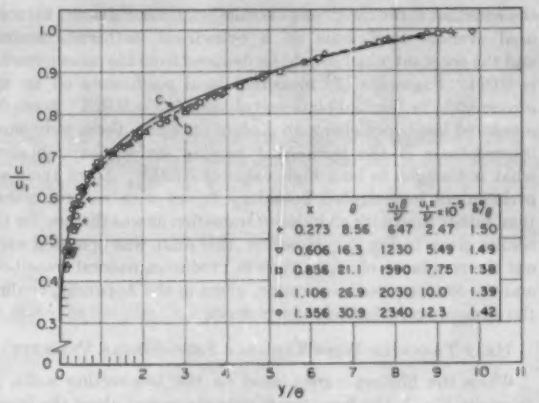


FIG. 4 BOUNDARY-LAYER VELOCITY DISTRIBUTIONS FOR A CONSTANT FREE-STREAM VELOCITY OF 165 FPS
(Curve b is distribution obtained by Klebanoff, reference 7, and curve c is the sixth-power law.)

Fig. 4, and compares them to values obtained from Equation [5] and also to values obtained from the relation

$$c_f = \frac{0.025}{\left(\frac{u_1\theta}{\nu}\right)^{1/4}} \quad [6]$$

where in this latter evaluation the values of the momentum thickness θ were obtained from the velocity distributions.

TABLE 2 VALUES FROM VELOCITY DISTRIBUTIONS OF FIG. 4

x, ft	0.273	0.606	0.856	1.106	1.356
c_f, vel dist	0.0054	0.0045	0.0043	0.0039	0.0037
c_f, Equation [5]	0.0049	0.0042	0.0039	0.0037	0.0036
c_f, Equation [6]	0.0050	0.0042	0.0040	0.0037	0.0036

The friction coefficients deduced from the velocity distribution are 3 to 9 per cent higher than the predicted values, while the predicted values are in close accord, since the experimental and predicted values of the momentum thickness θ differed by only 10 per cent. The mean speed distributions of Fig. 4 are themselves not known accurately enough for the resolution of errors of this magnitude, and it can only be concluded that for the speed of 167 fps, there is a close approach to the ideal "flat-plate" condition. This cannot be extrapolated, however, to all the speeds for which heat-transfer results are presented in Fig. 3, where the results for the lowest speed run reveal that the roughness was ineffective in promoting immediate transition.

The experimental results of Fig. 3 can be specified generally within ±5 per cent by the prediction of the von Karman analogy which, however, refers to an isothermal rather than a constant flux surface. Rubesin (3) predicts the heat-transfer coefficients for a constant flux surface to be 6 per cent higher than for an isothermal surface, so that the present results would indicate an "isothermal" correlation this much below the von Karman prediction. The uncertainties involved in the accuracy of the measured heat-transfer coefficient, combined with those related to the friction coefficient, make impossible this accurate a specification of the true isothermal value of the heat-transfer coefficient, as would be indicated by the present result. It is clear that the mean value of these results is definitely below the Colburn prediction and in this there is agreement with other experimental results.

Comparison to other experiments can be made best by noting that, if the results are expressed as

$$\frac{hx}{k} = C\left(\frac{u_1 x}{\nu}\right)^{4/5}$$

the constant C for the present results is 0.0236. Jakob (4) measured average coefficients on a cylindrical isothermal surface and the constant which would be deduced from his measurements is 0.024. Sugawara (5) measured local coefficients on an approximately isothermal plate and obtained $C = 0.023$. Scesa (6) measured local coefficients on a plate on which there were small interruptions in the isothermal heating surface and obtained what is thought to be a high value of 0.0255. All of these experimenters investigated boundary layers with natural transition, rather than with an induced transition as was the case for the results given in Fig. 3. However, this plate was operated without the roughness element so as to produce a natural transition, and the results of such operation, given in the Appendix, confirm the comparisons that have been made.

Heat Transfer With Variable Free-Stream Velocity

When the blisters were placed on the test-section walls, as shown in Fig. 1, the free-stream velocity varied along the length of the plate. A first relatively slight acceleration was followed by a strong one as the flow approached the blister, and there then followed a region of retarded flow to the end of the plate. Fig. 5 reveals the experimental results for the heat-transfer coefficient,

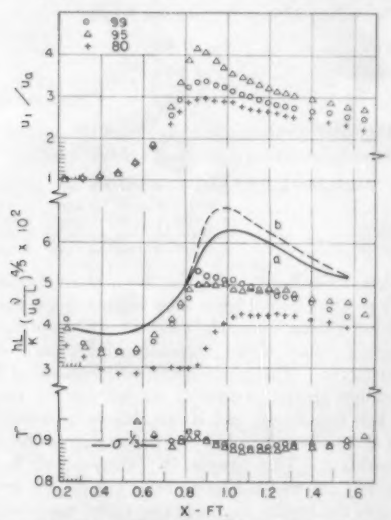


Fig. 5 Results With Turbulent Flow and Variable Free-Stream Velocity

(Curve a is prediction based on velocity distribution of Run 80, and curve b on distribution of Run 99. Physical properties contained in representation were evaluated at surface temperature and pressure at initial thermocouple location. Magnitudes of velocity u_a are indicated below, together with the factor by which ordinate of heat-transfer representation can be multiplied to obtain local heat-transfer coefficient in Btu/hr sq ft deg F.)

Run	u_a, fps	
95	215	1130
99	164	940
80	94	600

where one set of results is shown for each of three magnitudes of the velocity. Unlike Fig. 3, the heat-transfer coefficients are not corrected for unheated starting length because of the uncertainty of the correction in this case and the theory to which the results are ultimately to be compared. The representation used in Fig. 5, in which the heat-transfer coefficient is incorporated into the group

$$\frac{hL}{k}\left(\frac{\nu}{u_a L}\right)^{4/5}$$

confirms the theory by achieving a correlation of the results for the two highest-speed runs, though it fails for the low-speed run because the flow there was apparently not completely turbulent. This point is discussed in detail later.

Values of the recovery factor, shown in Fig. 5, are of the order of the cube root of the Prandtl number, and they do not appear to be affected by the substantial variation in free-stream velocity that occurred.

The rationalization of the heat-transfer performance exhibited in the case of variable free-stream velocity depends upon a solution of the hydrodynamic and thermal problem, which can be simplified in the first approximation by an assumption of constant properties, to render separate the hydrodynamic and thermal problems. Many methods have been proposed for the prediction of turbulent boundary-layer behavior with variable free-stream velocity, though most refer primarily to retarded flows. Here the simplest possible method was employed. by taking the parameter H as constant and using the friction law of Equation [6]. Then the integral form of the momentum equation was employed with an initial condition based on the measured momentum thickness at the first station. With $H = 1.28$, the following form of the momentum equation is obtained

$$\frac{d\theta^{4/5}}{dx} + \frac{5}{4}(3.28)\frac{\theta^{4/5}}{u_1}\frac{du_1}{dx} = \frac{5}{4}\frac{(0.0125)}{(u_1/\nu)^{1/5}} \quad \ldots\ldots [7]$$

Integration gives, for $\theta = 0$ at $x = 0$

$$\left(\frac{\theta}{L}\right)^{4/5}\left(\frac{u_a L}{\nu}\right)^{1/5} = \left(\frac{u_a}{u_1}\right)^{4.10}(0.0156)\int_0^{x/L}\left(\frac{u_1}{u_a}\right)^{3.55} d\left(\frac{x}{L}\right) \quad \ldots\ldots [8]$$

Fig. 6 contains results obtained from Equation [8] and the free-stream velocity distribution obtained for Run 80, the low-speed run. A calculation made with $H = 1.33$, typical of a $^1/_6$-power low-velocity distribution, produced almost the same result.

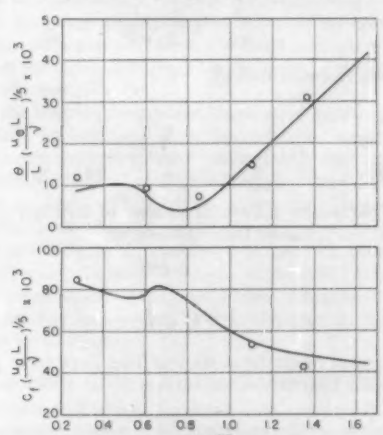


Fig. 6 Momentum Thickness and Friction Coefficient

(Momentum thickness and friction coefficient are presented in groups indicated by theory leading to Equation [8]. Curves are predictions made for velocity distribution of Run 80, and points are values obtained from boundary-layer velocity distributions obtained for conditions corresponding to Run 99, which are shown in Fig. 7.)

Fig. 7 reveals the velocity distributions in the boundary layer obtained at flow conditions corresponding to Run 99, at higher velocity, and the momentum thicknesses and friction coefficients obtained from these distributions have been placed in Fig. 6 for comparison with the predicted values. Good correspondence exists, even though the compressibility effects are sufficient in

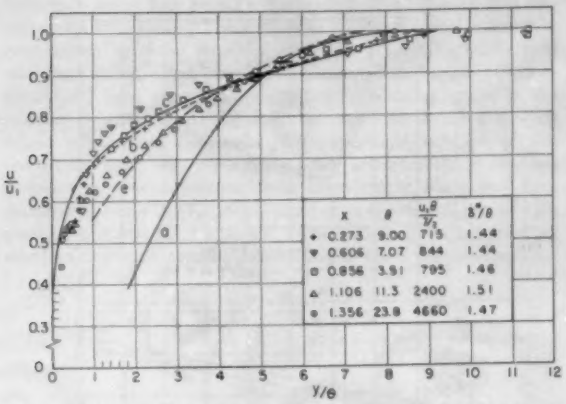


FIG. 7 VELOCITY DISTRIBUTION IN BOUNDARY LAYER FOR CONDITIONS CORRESPONDING TO RUN 99

(Curve: *a*, Blasius distribution; *b*, Klebanoff; *c*, sixth-power law; *e*, von Doenhoff, $H = 1.6$. (ν_c is the kinematic viscosity at the local pressure and stagnation temperature.)

Run 99 to cause some change in the free-stream-velocity distribution. These effects are not large, however, for integration of Equation [8] with the velocity distribution of Run 99 does not alter materially the result.

The agreement between the predicted and experimental values dealing with the flow indicates the hydrodynamic analysis to be a satisfactory basis for the evaluation of the heat transfer. One method of heat-transfer prediction is the application of the analogy as expressed by Equation [3] or [4] directly to the friction coefficients as shown in Fig. 6. This fails conspicuously, yielding heat-transfer coefficients which are high, particularly in the acceleration region, where the momentum thickness becomes small. There too the momentum-thickness Reynolds number is almost constant and the heat-transfer conditions approach, in a sense, those of pipe flow, in which the thermal history is important and is accounted for by the use of the mean fluid temperature in the expression of the effective temperature difference. While such an analysis is possible and at least partially successful in the case under consideration, it lacks precision because of the impossibility of defining exactly the mean temperature of a boundary-layer flow. The method employed instead is also approximate but has the advantage of an appeal to established theories concerning the heat transfer to boundary layers.

The effect of the thermal history of the flow on the heat transfer is based upon the result obtained by Rubesin's (3) analysis for the heat transfer to a constant velocity flow for which the surface temperature exceeds the free-stream temperature by $t + \Delta t$ in the region $0 < x < a$ and by t for $x > a$. Then the heat flux at a point $x > a$ is

$$q = h_0 (t + \Delta t) - h_0 \Delta t \left[1 - \left(\frac{a}{x} \right)^{39/40} \right]^{-7/8} \quad \ldots\ldots [9]$$

where h_0 is the local heat transfer for a completely isothermal surface. If the exponent 39/40 is taken as approximately unity, and the exponent 7/39 as approximately 1/5, the total heat transfer in the length x is found to be

$$Q = \frac{5}{4} h_0 x \left[t + \Delta t - \Delta t \left(\frac{x - a}{x} \right)^{4/5} \right] \quad \ldots\ldots [10]$$

For x much larger than a this can be approximated again to give

$$Q = \frac{5}{4} h_0 x t \left[1 + \frac{4}{5} \frac{a \Delta t}{x t} \right] \quad \ldots\ldots [11]$$

Similar assumptions give the following approximation for Equation [9]

$$q = h_0 t \left[1 - \frac{1}{5} \frac{a \Delta t}{x t} \right] \quad \ldots\ldots [12]$$

Group $(a\Delta t)/(5xt)$ indicates the effect of a departure from a completely isothermal history, and its elimination between Equations [11] and [12] gives

$$\frac{q}{t} = \frac{h_0}{\left[0.80 + \frac{4Q}{25\, qx} \right]} \quad \ldots\ldots [13]$$

Now on the flat plate, for which this analysis applies

$$x = \frac{8}{5} \frac{\theta}{c_f}$$

and the relation of the local coefficient to the isothermal coefficient becomes

$$\frac{h}{h_0} = \frac{1}{\left[0.80 + \frac{Q}{10\, q} \frac{c_f}{\theta} \right]} \quad \ldots\ldots [14]$$

This result, still substantially correct for a sectionally isothermal surface, can be arbitrarily extended to the case of constant heat flux, for which $Q = qx$. Equation [14] then indicates that $h = 1.04\, h_0$, which is 2 per cent below Rubesin's value of $1.06\, h_0$. Finally, the result for constant heat rate is applied arbitrarily to the case of constant heat flux with variable free-stream velocity. With x_n as the heated length, Equation [14] is then taken in the form

$$\frac{h}{h_0} = \frac{1}{\left[0.80 + \frac{x_n c_f}{10\, \theta} \right]} \quad \ldots\ldots [15]$$

A prediction of the heat-transfer coefficient has been made by combining the local friction coefficient from Fig. 6, the Colburn analogy, Equation [3], and Equation [15], and this is shown as curve *a* in Fig. 5. A similar prediction, made for the velocity distribution of Run 99, is shown also and it departs by a maximum of 10 per cent from curve *a*. While the local friction coefficient is but slightly affected by the change in velocity distribution, the heat-transfer coefficient, because of its more direct dependence on velocity as exhibited in the Stanton number, changes appreciably. Previous experience is repeated in the best agreement between experiment and the prediction made for the velocity distribution in closest approach to incompressible flow. This agreement, shown in Fig. 5, would be improved by following the dictate of the results of Fig. 3 and using Equation [4] in the prediction method rather than Equation [3]. This would reduce the predicted heat-transfer coefficients by about 10 per cent, and secure almost complete agreement between theory and experiment up to the point of flow deceleration. In view of the approximate nature of the theory leading to Equation [15], the agreement is remarkable.

The less satisfactory agreement in the deceleration region points to the inadequacy of Equation [3] (or Equation [4]) in the region of decelerated flow, since the contribution of Equation [15] is negligible in this region where the momentum thickness is large and increases rapidly. The 15 per cent difference (based on Equation [4]) existing in this region has not yet been explained.

Fig. 8 is a particular interpretation for Run 80, for which the correlation in Fig. 5 failed. The results for Run 79, at an even

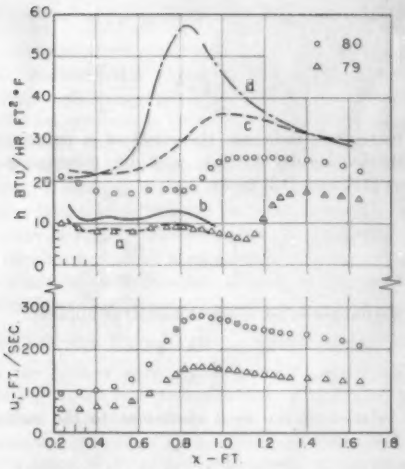


FIG. 8 HEAT-TRANSFER COEFFICIENTS AT LOW VELOCITIES

(Curves *a* and *b* are laminar-flow predictions for Runs 79 and 80; curve *c* is turbulent-flow prediction for Run 80. Curve *d* is value of h_0 to which Equation [15] is applied in order to obtain curve *c*.)

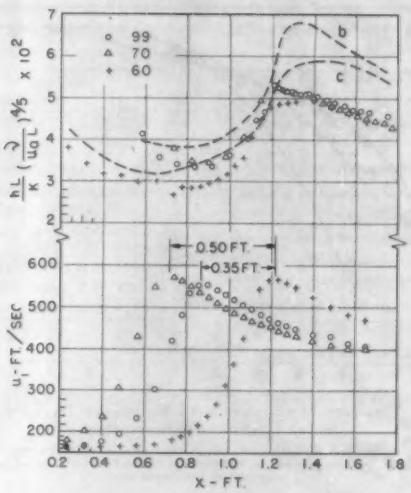


FIG. 9 EFFECT OF CHANGING THE BLISTER POSITION

(Curve *b* is the same as curve *b* of Fig. 5, and curve *c* is a similar prediction for Run 60. The velocity distributions are shown in their true positions but the heat-transfer results for Runs 99 and 70 have been shifted back by amounts shown in the figure. Factors by which ordinate can be multiplied to obtain local heat-transfer coefficient are: Run 99, 940; Run 70, 936; Run 60, 937.)

lower velocity, are also shown, together with the laminar-flow prediction for it. These demonstrate the complete ineffectiveness of the grit roughness at low speeds, the boundary layer for Run 79 being laminar until transition ultimately occurs in the region of retarded flow. Both the laminar and turbulent-flow predictions are shown for Run 80 and the actual heat-transfer coefficient is situated, remarkably, between the two. The character of the flow is not known, though mean speed distributions taken in the boundary layer appear to be like those of turbulent flow.

Fig. 8 contains also a curve showing the result of the combination of the local friction coefficient and Equation [3], and, by contrast with curve *c*, demonstrates the contribution of Equation [15] to the magnitude of the predicted heat-transfer coefficient.

Of corollary interest are results obtained by changing the position of the blisters in the tunnel, so that the boundary-layer thickness differs at the points where the acceleration becomes strong. Its strength is such that the boundary-layer thickness quickly becomes independent of the initial thickness, and in the region of maximum velocity the only effect of different blister positions is a difference in the amount of heat that has been transferred to the layer. This should effect a reduction in the heat transfer in the region of maximum velocity as the upstream heated length is increased. Fig. 9 reveals the effect in the experimental results which have been displaced so that the points of maximum velocity coincide. The predictions shown there were made for the higher velocity (Run 99) and are therefore slightly higher than the best interpretation, but they too reveal the relative effect, though to a degree greater than that shown experimentally.

CONCLUSIONS

This experimental study has considered the heat transfer to turbulent boundary layers with variable free-stream velocity, under circumstances in which a boundary layer established under relatively constant velocity conditions was subjected to a strong acceleration and a subsequent deceleration of the flow. The heat-transfer coefficients were affected substantially by the acceleration.

Methods of prediction based on the assumption of incompressible flow give the correct heat-transfer performance provided an account is made of the thermal history of the flow. The method used in this paper, while approximate, produced excellent results in the acceleration region and fair though not completely satisfactory results in the region of retarded flow.

ACKNOWLEDGMENTS

The results presented were obtained in heat-transfer research at the University of California sponsored by the United States Air Force under Contract AF 33(616)-348. The authors also acknowledge the contribution of H. W. Chan, who obtained many of the data and calculated many of the results.

BIBLIOGRAPHY

1 "Forced Convection From Nonisothermal Surfaces," by J. Klein and M. Tribus, Heat Transfer: A Symposium, University of Michigan Press, Ann Arbor, Mich., 1953.

2 "Turbulent Boundary Layers in Adverse Pressure Gradients," by F. G. Clauser, *Journal of the Aeronautical Sciences*, vol. 21, February, 1954, pp. 91–108.

3 "The Effect of an Arbitrary Surface Temperature Variation Along a Flat Plate on the Convective Heat Transfer in an Incompressible Turbulent Boundary Layer," by M. W. Rubesin, National Advisory Committee for Aeronautics, TN 2345, 1951.

4 "The Heat Transfer From a Cylindrical Surface to Air in Parallel Flow With and Without Starting Sections," by M. Jakob and W. Dow, Trans. ASME, vol. 68, 1946, pp. 123–134.

5 "Heat Transfer on the Surface of a Flat Plate in the Forced Flow," by S. Sugawara and T. Sato, Mem. Fac. of Engineering, Kyoto University, Japan, vol. 14, no. 1, 1952.

6 "An Experimental Investigation of Heat Transfer to Air From a Flat Plate With a Stepwise Discontinuous Surface Temperature," by S. Seesa and F. Sauer, Trans. ASME, vol. 74, 1952, pp. 1251–1255.

7 "Some Features of Artificially Thickened Fully Developed Turbulent Boundary Layers With Zero Pressure Gradient," by P. Klebanoff and Z. Diehl, National Advisory Committee for Aeronautics, TN 2475, 1951.

Appendix A

HEAT TRANSFER WITH CONSTANT VELOCITY AND NATURAL TRANSITION

The plate also was operated with constant free-stream velocity and without the roughness element at the nose, so that the

boundary layer was initially laminar, with turbulent flow obtained by transition. The results so obtained are of interest in comparison to those obtained with the artificially established turbulent flows, which have been presented in Fig. 3. Corrections have been applied for an unheated starting length in the manner described, with laminar-flow corrections being applied where appropriate. Because of the variable location of the transition, there was necessarily some arbitrariness in the application of the correction, with the laminar correction used completely when the Reynolds number at the first thermocouple position was less than 1×10^6. This arbitrariness contributed some scatter to the results, particularly in the transition region.

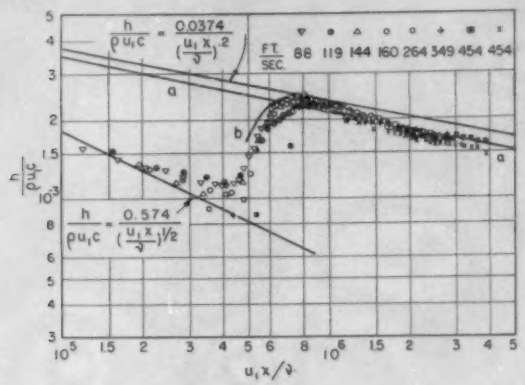


FIG. 10 STANTON NUMBERS FOR CONSTANT VELOCITY WITH NATURAL TRANSITION

(Curves are same as those of Fig. 3. Curve *b* is discussed in text. All physical properties involved in representation of experimental results were evaluated at surface temperature and pressure.)

Fig. 10 reveals the results for the Stanton number, and is distinguished by the relatively early transition, occurring at a Reynolds number of 4.5×10^5. This was confirmed by mean speed distributions in the boundary layer, which approximated the Blasius profile for Reynolds numbers less than 4×10^5, and typical turbulent profiles for Reynolds numbers greater than 7×10^5. The cause of this early transition, despite the relatively low turbulence level of the tunnel, was not established.

The agreement of the Stanton numbers in the range of turbulent flow with the prediction of the von Karman analogy is equivalent to that obtained with artificially induced turbulent flow. Although this is expected, some explanation is required, for the local friction coefficients, at least immediately after transition, differ substantially from values that exist at the same location if the boundary layer is turbulent from the leading edge. For example, with a free-stream velocity of 165 fps, the mean velocity distributions in the turbulent part of the boundary layer, as obtained at tranverse stations 3, 4, and 5, yielded the results contained in Table 3.

TABLE 3

x, ft	0.856	1.106	1.356
θ mils	10.6	16.6	21.6
c_f, vel dist		0.0046	0.0042
c_f, Equation [5]		0.0037	0.0036
c_f, Equation [6]		0.0041	0.0039

Extrapolation of the values of the momentum thickness indicates them to characterize a turbulent boundary layer which would be initiated at $x = 0.53$ ft. The comparison of values of the friction coefficient deduced from the velocity distributions and those given by Equation [6] indicates that even on this basis the behavior of the turbulent layer is not explained completely.

The analysis used for the case of variable velocity, containing Equation [15] to account for the thermal-energy content of the layer, was applied to these results, with the effective initial point of the turbulent layer at a Reynolds number of 4.5×10^5, corresponding to the effective initial point deduced from the observations made at 165 fps. Curve *b* of Fig. 10 is the result of the application of this method. It shows that the trend of the heat-transfer coefficient in the latter part of the transition region and the rapid approach to the analogy prediction for a completely turbulent flow can be explained in terms of the effect of the thermal-energy content of the layer.

The prediction shown as curve *b* applies only to a velocity of the order of 165 fps because of the fixed length of unheated starting length on the model. If the prediction is made for higher velocities, the curve rises slightly above the Colburn line before becoming asymptotic to it, while for lower velocities the approach is even more gradual than shown by the curve in Fig. 10. The results do not show such an effect in detail, and the effect is so small as to be within the experimental inaccuracy.